DOCUMENT RESUME

ED 102 200 95 TM 004 257

AUTHOR	Andrews, Theodore E., Ed.
TITLE	Assessment.
INSTITUTION	New York State Education Dept., Albany. Div. of Teacher Education and Certification.; State Univ. of New York, Albany.
SPONS AGENCY	National Center for Improvement of Educational Systems (DHEW/OE), Washington, D. C.; Office of Education (DHEW), Washington, D.C. Teacher Corps.
PUB DATE	74
NOTE	130p.
AVAILABLE FROM	National Dissemination Center for Performance Education, Syracuse University, Syracuse, New York 13210
EDRS PRICE	MF-$0.76 HC-$6.97 PLUS POSTAGE
DESCRIPTORS	Certification; Costs; *Educational Assessment; Educational Research; Inservice Teacher Education; Lesson Plans; *Measurement Techniques; Models; Objectives; *Performance Based Teacher Education; Performance Criteria; Preservice Education; Program Evaluation; State Departments of Education; Student Evaluation; *Teacher Evaluation; Teaching Methods

ABSTRACT

"The Role of the State in Performance-Based Teacher Education-Certification" by Robert Roth creates a context for viewing how state agencies are approaching performance education. Peter Airasian then explores the value questions that are at the heart of evaluation issues. Fred McDonald looks at "The State of the Art in Performance Assessment," and Barak Rosenshine lists recommendations concerning the research dilemmas. James Popham, in three papers, touches on the problems of selecting assessment systems, developing performance tests, and identifying minimal competencies. Del Schalock then details what occurs when "Moving from Conceptualization to Practice in Assessment." Concerning costs and teacher concerns, two educators, Bruce Joyce and Herbert Hite present their conclusions. Beatrice Ward discusses the cost factors involved in developmental work at the Far West Regional Laboratory. Finally two teachers, Sandra Feldman and Bernard McKenna, note their interest in the potential of performance-based teacher education and reveal their most serious concerns. (RC)

BEST COPY AVAILABLE

PERFORMANCE EDUCATION

ASSESSMENT

THE UNIVERSITY OF THE STATE OF NEW YORK
THE STATE EDUCATION DEPARTMENT
DIVISION OF TEACHER EDUCATION AND CERTIFICATION
and
MULTI-STATE CONSORTIUM ON PERFORMANCE-BASED
TEACHER EDUCATION

TM004 257 — ED102200

ED102200

ASSESSMENT

PETER AIRASIAN • SANDRA FELDMAN • JESSE H. GARRISON • HERBERT HITE
BRUCE R. JOYCE • BERT Y. KERSH • FREDERICK McDONALD • BERNARD McKENNA
W. JAMES POPHAM • SARAH ROSENSHINE • ROBERT A. ROTH
H. DEL SCHALOCK • BEATRICE A. WARD

Theodore E. Andrews, Editor

MULTI-STATE CONSORTIUM ON PERFORMANCE BASED
TEACHER EDUCATION

CONTRIBUTORS

Peter W. Airasian
Boston College
Boston, Massachusetts

Sandra Feldman
United Federation of Teachers
New York City, New York

Anne M. Gardner
Oregon College of Education
Monmouth, Oregon

Herbert Hite
Western Washington State College
Bellingham, Washington

Bruce R. Joyce
Teachers College
Columbia University
New York City, New York

Burt S. Kaufman
Oregon College of Education
Monmouth, Oregon

Frederick J. McDonald
Educational Testing Service
Princeton, New Jersey

Dorward H. McKenna
National Education Association
Washington, D.C.

M. James Popham
University of California, Los Angeles
and
The Instructional Objectives Exchange

Robert L. Rath
Coordinator of Performance Education Project
New Jersey Department of Education

Ruth Ravenstein
Bureau of Educational Research
University of Illinois at Urbana

H. Del Schalock
Teaching Research
A Division of the Oregon State System of
Higher Education
Monmouth, Oregon

Gunnar L. Ward
Associate Director, Teacher Education
Far West Laboratory for Educational
Research and Development

Foreword

Introduction

The Role of the State in Performance-Based Teacher Education/Certification

by

[redacted]

Certification and Education

The movement toward development of performance-based teacher education and certification programs has taken many years, growing slowly in the early years. Over the past several years, however, the movement has grown rapidly, and it is now the subject of a great deal of attention. Many states are currently involved in the study of performance-based programs for teacher education and certification. The purpose of this paper is to discuss the role that the state can play in promoting performance-based programs. The discussion will center on three areas: (1) the relationship between the state and the profession, (2) the relationship between the state and the institutions of higher education, and (3) the relationship between the state and the local school district. Each of these areas will be discussed in relation to the role that the state can play in promoting performance-based programs.

Before we begin, it is important to recognize the fundamental differences between the purposes of teacher education and those of certification. Generally speaking, teacher education programs are designed to prepare individuals for entry into the profession, whereas certification programs are designed to assure the public that individuals who have met certain standards are permitted to practice their profession. Thus, while the two functions are closely related, they are not identical. Teacher education programs are concerned with the development of professional competence, while certification programs are concerned with the assessment of professional competence. Both functions are important, but they are not the same.

It is interesting to note that more than seventy-five percent of the states have enacted performance-based certification laws or regulations. Many states have also mandated performance-based teacher education programs. As a result of the enactment of the Education Amendments of 1976, many states are required to develop performance-based programs. The purpose of this paper is to discuss the role that the state can play in promoting performance-based programs. The discussion will center on three areas: (1) the relationship between the state and the profession, (2) the relationship between the state and the institutions of higher education, and (3) the relationship between the state and the local school district. Each of these areas will be discussed in relation to the role that the state can play in promoting performance-based programs.

Before we begin, it is important to recognize the fundamental differences between the purposes of teacher education and those of certification. Generally speaking, teacher education programs are designed to prepare individuals for entry into the profession, whereas certification programs are designed to assure the public that individuals who have met certain standards are permitted to practice their profession. Thus, while the two functions are closely related, they are not identical. Teacher education programs are concerned with the development of professional competence, while certification programs are concerned with the assessment of professional competence. Both functions are important, but they are not the same.

[redacted]

The difference between certification and teacher education is more significant when considering state responsibility for assessing certification. While there are many ways to assess performance-based certification, one of the most common is through the assessment of competencies. The assessment of competencies is a process that is used to determine whether or not an individual has the knowledge, skills, and abilities that are necessary for successful performance in a particular role. The assessment of competencies is a process that is used to determine whether or not an individual has the knowledge, skills, and abilities that are necessary for successful performance in a particular role.

In the area of certification, responsibility for assessing competencies is the responsibility of the state. The state is responsible for the assessment of competencies in the areas of [redacted].

Local schools, [redacted] are responsible for the assessment of competencies in the areas of [redacted]. The assessment of competencies in the areas of [redacted] is the responsibility of the local school district. The assessment of competencies in the areas of [redacted] is the responsibility of the local school district.

[redacted]

dividual, there is the possibility, depending on the program, for freedom to define goals, for flexibility in personality, method, and philosophy. Reflecting this viewpoint, William Drummond, a former associate in the Washington State Department of Education, urged that "State departments of education, therefore, should foster creativity and intellectual freedom and promote programs of teacher education which support and cherish uniqueness and individualism." [20]

The Washington model, therefore, also rejects the regulatory role of a state department of education. Wendell Allen, as Washington's assistant superintendent of public instruction, concluded

> To emphasize this regulatory role is to protect the status quo. When the rule is the thing, change must come before there can be a new rule. There is danger in this circumstance that the major energies of the agency will be spent on administrative rather than leadership functions.[21]

An essential point to note is the prevalence of multiple standards, lack of uniformity, less legal need for an empirical base, and no single set of standards. Should all of the above factors be deemed advisable, then a particular state might select this model.

New York has envisioned a very similar type of program.[22] Four process standards have been established to be utilized for the development of pilot projects. The standards require the establishment of a policy board made up of representatives of teachers, school districts, colleges, and teacher education students. This group considers the objectives of the schools involved, the competencies teachers need to be successful in that environment, as well as those qualities desirable for all teachers, and acceptable evidence for attainment of competencies. The policy board then will establish individualized programs for the preparation of teachers to meet these criteria. Finally, a management system must be established. Trial projects may be designed for initial or continuing certification or both. The State Department of Education will exercise its legal responsibility for program approval. Note the decentralized role and the belief that performance criteria are mostly situation specific.

Vermont [23] has expanded the decision-making base to local school districts. A local school district may develop a program for the inservice training and professional advancement of its staff and may apply to the State Department of Education for approval to recommend issuance and renewal of all certificates at the local level. The appropriate certificate will be issued by the State Department of Education.

The local district must submit evidence that the teachers, school board, and administrative personnel have participated in the planning and development of the program. The local program must include provision for job description, task analysis, and performance criteria for all education personnel. An approved program approach is in effect for college teacher preparation programs.

Washington, New York, and Vermont are case studies that fall into the process model. Local decision making characterizes these attempts, assuming what is acceptable in one situation may be unacceptable in another.

Moving slightly along the continuum away from the process model but within the local decision-making framework, there is a model suggested by Lierheimer [24] which we may call the informational model. The central thesis of Lierheimer's proposal is that the state's role is not to make judgments but to maintain records. He suggests the students be tested over a multitude of factors including actual teaching performance. There is a possibility here for utilization of performance criteria, but the testing is not done by the state.

Decentralization is emphasized in this approach with local school teams conducting the evaluation of the competence of potential teachers. Ultimately, the agency to decide on teacher performance for licensing purposes would be the school. The function of the state is to monitor the local evaluation but not impose state standards. Although evaluation systems would be approved by the state there would be no uniform techniques for verification of classroom performance. The state office would maintain a data bank on all teaching personnel in the state.

[20] William H. Drummond, "Conference Commentary" in The Seattle Conference: The Role of the State Department of Education in Teacher Education, Olympia, Washington: State Superintendent of Public Instruction, 1967, p. 74.

[21] Wendell C. Allen, "State Government and Teacher Education—A Different Role for the State Education Agency," in The Seattle Conference, p. 78.

[22] New York State Education Department, "A New Style of Certification," Albany, New York: New York State Department of Education, Division of Teacher Education, March 15, 1971.

[23] Vermont State Department of Education, "Regulations Governing the Certification of Educational Personnel," Montpelier: Vermont State Board of Education, Department of Education, July 1, 1971.

[24] Lierheimer, "Give Up The Ship."

A unique feature of this model is that the state accumulates information on an individual but makes no decision in reference to competence. The major role of the state is to provide resources. The local district is provided with the information, and it is at this level that decisions are made as to whether the individual's competence fits the particular situation. The underlying assumption is that values and competencies are situation specific and hence require local evaluation. Currently, there are no states utilizing this informational model. Again, analysis of the model should be made in terms of all the issues identified earlier.

This model can be modified to interject more state control and greater uniformity. Minimum standards could be set by the state for the various competencies or groups of competencies. These minimum standards would be established for certification purposes. The state would still maintain its individual data bank and local districts could use the information for hiring purposes. This modified model would be farther along the continuum in terms of state control and decision making.

Another open-type model which does not provide quite as broad a decision-making base is being developed by the state of Florida. In this case consortia are not designed for purposes of initial certification although inservice programs are developed by local districts. This "facilitation model" utilizes the college approved program approach commonly in practice among the states.

The program approval regulations are somewhat process in nature indicating prescribed activities, but they are content standards as well, identifying courses necessary for certification. There are alternatives to the content regulations which provide for performance-based programs.

> An institution may, instead, specify the competencies which its graduates will be expected to demonstrate, identify the procedures by which those competencies will be measured, and then develop a program which leads to those competencies. Once such a program is approved, its graduates will receive regular teaching certificates with no penalties. Institutions are now being encouraged to develop competency-based programs.[25]

In this model control is in the colleges, but direction is provided by the state. The colleges develop their own competencies, but these are consistent with state course requirements. There is additional direction and stimulus provided by the state, however, which facilitates development of such programs. The state is compiling a catalog of teaching competencies which will eventually be validated through research. These competencies, or performance criteria, will be provided to the colleges to facilitate their program development. These particular criteria, however, will not necessarily be mandated and certainly all will not be required of a given institution. Other facilitating procedures by the State are assembling of training materials based on performance criteria and staff development for teacher trainers. The emphasis is on facilitation. Decision making is somewhat diffused but the role of the state is stronger than in previous models. The faciiitation model presents different responses to the issues.

The remaining two models to be discussed can be grouped under a heading of central decision making. The first two models, you may recall, were local decision-making types, with the facilitation model being somewhere between. These last two models are at the other extreme end of the continuum.

One approach to performance-based certification is to establish performance criteria at the state level. This approach supports a strong state role and a uniform set of standards. It guarantees that each certificated individual has at least a minimum set of competencies. These criteria could be utilized as a state test or part of an approved program. The focus here, however, is on the approved program approach.

The manner in which these criteria are stated significantly affects the impact they will have on teacher education programs and the role of the state. The performance criteria can be stated in generic terms which then serve as guidelines for further specification by teacher preparation institutions. This guidelines model increases centralized authority yet does provide a certain degree of participation on the part of the colleges or consortia.

As an example of competencies consistent with the guidelines model, it might be required that the teacher candidate demonstrate the ability to diagnose areas of student deficiency, maintain a classroom environment which motivates students to learn, plan an instructional unit, employ a variety of instructional techniques, etc.

Utah[26] recently adopted at the state level a set of

[25] Florida Department of Education, "The Florida Program for Improving the Training, Evaluation, and Licensure of Educational Personnel," draft number 2. Tallahassee: Educational Research and Development Section, April 7, 1971.

[26] Utah State Board of Education," Recommended Proficiency Guidelines for Media Endorsements," Salt Lake City: Utah State Board of Education, 1972.

BEST COPY AVAILABLE

Robert A. Roth

performance criteria for instructional media, some of which approximate the guidelines model type of criteria. Prerequisite to a Basic Media Endorsement are a bachelor's degree and a teaching certificate. An examination for proficiency conducted by a recommending institution (with an approved certification program) is then administered. The recommending institution is free to determine how the competency will be demonstrated or ascertained, but a candidate may request an opportunity to demonstrate a competency whenever he feels he is ready. Competencies may be demonstrated one at a time. Candidates who perform satisfactorily will be considered as having met the endorsement requirement regardless of the route taken to obtain the competency.

Proficiency must be demonstrated in five areas. Some examples of performance standards are as follows:

—Using media selection tools of his choice, the candidate will identify the tools he has selected and include a rationale for the choice of each.

—The candidate will explain what one would do to select new subject headings for materials which are not considered in "Sears List of Subject Headings."

—The candidate will demonstrate proficiency in mounting pictures by producing one acceptable example of the following:

(1) dry mount on a hard surface, using dry mounting tissue

(2) dry mount, using dry mounting cloth

(3) rubber cement mount

(4) laminate with thermo copy machine, adhesive acetate, or heat press.

At the extreme end of the continuum we have what can be termed the "prescriptive model." In this system the state provides very specific performance criteria (behavioral objectives) which are utilized by the colleges as objectives and evaluative criteria. This is the most dominant of the state roles within an approved program approach with an emphasis on the administrative and regulatory function of a state education agency. Uniformity in certification with a single set of standards is the essential feature.

The state of New Jersey is currently studying the feasibility of such a performance-based certification system. Specific performance criteria are being developed for use on a statewide basis as certification standards. How specific these will be has not as yet been determined. There are two unique aspects to the New Jersey approach, however, that broaden the base of decision making. The performance criteria are being developed by task forces composed of a crosssection of educators from across the state, representing teachers, administrators, college students, college professors, the State Department, and the various professional associations. These criteria, therefore, are not developed by the State Department but represent a consensus of professional educators in the state. In addition, evaluation of prospective teachers may involve schools, colleges, and professional associations, a resemblance to the consortium idea.

Clearly, the statewide involvement in development of criteria adds considerable power to the approach. It presents a decided advantage over development of criteria by a state department or even a college or university. It appears to have greater validity and is more likely to find statewide acceptance. Significantly, "it has been generally agreed that whoever determines certification requirements controls the program of preparation." [27] Thus, in this instance, control is more in the hands of the total profession.

The guidelines model can be developed by the same method. The difference between the two models then lies in the specificity of the criteria. How does this difference relate to the issues, and how do these two models compare with the open end of the continuum in terms of the issues?

[27] **Improving State Leadership, p. 5.**

In the process model, teacher preparation institutions have maximum curricular freedom. The guidelines model allows the institutions the opportunity to develop the specific performance criteria while the prescriptive model does not provide for this. A comparison of performance criteria with traditional course list standards may be of value at this point. A course in tests and measurement is a familiar requirement in the course list system. The guidelines model would require competencies that are somewhat more specific, such as ability to evaluate student performance and ability to develop tests. The prescriptive model, however, would list a number of specific performances such as ability to formulate essay (multiple choice, etc.) test items and analyze tests for validity. Also, the evidence accepted that the performance had been achieved would be provided. Continuing our comparison, if used in a course list system, a prescriptive model would list the things that should be taught in a tests and measurement course rather than leaving this to the college.

Andrews has stated that "a required set of performance criteria could be just as moribund as rigid course requirements have been in the past."[1] It appears that the more specific the criteria the less freedom that exists. Recall that the approved program approach works "within the framework of generally agreed upon goals."[2] The possibilities for creativity through innovative programs can be achieved in the design of means to achieve the objectives, but not through alternative objectives. Two basic questions are at hand. First, is curricular freedom seen as being of value; and second, does a prescribed set of specific performance criteria significantly limit this freedom? A related question is whether or not the guidelines model offers a great deal more freedom than the prescriptive model.

A concern similar to the question of freedom is diversity. The process model allows, and even encourages, diversity among programs. Those in favor of diversity argue that there are varied philosophies of education requiring different teaching models. Any set of performance criteria is based on a theory of teaching and the teaching-learning process. Although not always articulated, the purposes of teaching are inherent in the criteria.

In the process model several teaching philosophies exist simultaneously with validation and development being ongoing processes. A set of specific criteria, however, relies on one teaching model and also establishes a particular value system. The problem here is that there is no empirical base to lead us to the correct model. As noted earlier, lack of an empirical base is a primary concern with performance-based certification. With a variety of program types, it can be argued, we recognize the developmental state of our knowledge base, whereas a single model seems a finality and demands empirical validation before being adopted. This accounts for the support of performance-based teacher education instead of certification.

Another point made by those favoring diversity is that performance criteria are situation specific. There are numerous contexts for teaching, both in terms of environment and educational philosophy. This requires different sets of competencies, at least in terms of the general situations (not for every school, etc.). There may not be enough in common to establish at least a minimum core of competencies at the state level. Washington, Vermont, and New York appear to believe in this as evidenced by their process models.

All of the above factors suggest multiple standards and diversity of programs. The initial question is whether these are valid concerns. The other position argues for more standardization and quality assurance. Inequities among programs are diminished. Certainly, the prescriptive model adheres to the latter viewpoint. The guidelines model does provide a certain degree of variability in that each institution can define the specific criteria to fit its needs. The prescriptive model insists on a single standard; the guidelines model offers some degree of multiple standards although minute when compared to the process model.

A frequent criticism of competency-based programs is the problem of writing performance criteria in the affective domain. This problem becomes amplified as we move across the continuum toward the prescriptive model. As an example, the guidelines model might require competence in developing teacher-student rapport. Each teacher preparation institution would be provided the freedom to determine not only how this might be developed but how it might be judged to exist. The prescriptive model, however, would specify the performance criteria necessary to achieve this, such as "uses student names," or "smiles or acknowledges student responses by nodding." The question is whether or not such criteria can be written on a statewide level. Ignoring the affective domain and concentrating on the cognitive and psychomotor would not be a viable alternative.

The reader may recall the issues raised concerning

[1] Theodore Andrews, **New Directions in Certification**, Washington, D.C.: Association of Teacher Educators, 1971, p. 10.

[2] **Improving State Leadership**, p. 7.

the rights of the individual as suggested by McDonald.[7] Are there opportunities for flexibility in personality, method, and philosophy? What about the right of the individual to define his own goals? Buckley and Miller, as members of the Pennsylvania State Department of Education, stated that

> Individual differences are not taken into account in blanket certification standards. We are convinced that the improvement of teacher preparation must take place at the point of initial preparation . . . with attention directed to individual needs within the context of general certification requirements.[8]

The process model provides for individual flexibility, and there are functioning programs which operate on these premises. The prescriptive model precludes much of this, at least in terms of the specific criteria required by the state. The individual does not have the freedom to define his own goals but he may have the opportunity to select his own method of achieving the objectives. At best, those favoring a uniform set of standards would find individual selection of goals to be undesirable and detrimental to certification.

The guidance model may provide a certain degree of individual choice but within the boundaries defined at the state level. The general objectives must be accomplished, but the specifics can vary with the individual. The manner in which one wishes to develop teacher-student rapport or plan for a lesson can vary significantly from another individual's method. The basic question is not just one of uniform standards versus flexibility, but the degree of each that is desirable.

Alternatives to Approved Programs

The discussion of issues and alternatives has thus far been limited to the approved program approach to state certification. Approved programs referred to those developed by colleges alone or by consortia. The evidence presented earlier in this paper suggest that approving programs is the most viable approach to performance-based teacher certification, and some specific criticisms of the state testing approach were described.

There are some teaching areas, however, that find themselves less rigidly tied to college preparation programs and thus are more amenable to alternative approaches. The area of vocational education, for example, is somewhat unique in that it usually relies on experienced professionals in the various trades to enter the teaching profession. There are other areas such as music, that also require specific skills unique to the particular profession. Educational fields such as these warrant consideration of alternative approaches that are not necessarily bound by college degree programs. These different approaches are not necessarily limited in application to the special teaching areas mentioned, however, as the alternatives may be utilized for any teaching field if desired.

A commonly discussed alternative to teacher certification is the establishment of a state testing procedure. There are several ways in which this can be implemented, some of which will be described here. A major argument against this approach (which was pointed out earlier) is that there exists no empirical base on which to construct a valid testing technique, particularly in view of varied teaching contexts. The predictive validity of any such examination device would have to be established.

It is again important to consider how the state testing models reflect the various issues. Questions of curricular freedom, individual freedom, and varied teaching philosophies should not be forgotten. The state testing approach to certification offers radically different responses to the issues when compared to the models within the approved program approach.

The informational model suggested by Lichtenberg can easily be modified to fit a state testing procedure. A set of behaviorally stated competencies could be formulated as certification descriptors. A teacher candidate's degree of accomplishment of each of the criteria could be indicated to form his competency profile. Minimum standards established for certification could be set by the state for each criterion or group of criteria. A system could be established (total score, weighted scores, etc.) to determine the individual's eligibility for certification. The state would still maintain its individual data bank and local districts could use the information for hiring purposes.

An important modification of the Lichtenberg informational model is that not only are minimum levels established for certification, but the testing of the candidate to determine his achievement of each criterion is done by the state, not through an approved program approach. The control of standards and verification of accomplishment reside in the hands of the state.

The modified informational model is but one variation of the state testing concept. Any outside agency or group of evaluators could be designated by the state to

---

[7] McDonald, "Philosophical Problems."

[8] F. R. Buckley and Norman Miller, "Broad Policy Concerns and Direction for a State Department of Education in Teacher Education," in The Seattle Conference, p. 65.

carry out the testing function. There is an opportunity to involve members of the profession in both development of criteria and actual use evaluating levels of teams who certify individuals. Instead of a profes. certification of minimum competency might be all that is necessary. Different certificates could be issued at different degrees of accomplishment or even different types of criteria. Evaluating boards or teams could again be used throughout the entire process.

It is generally assumed that the evaluation for certification would be done in a live classroom situation. An alternative would be to establish testing centers where specific skills would be evaluated such as flow board or microteaching. This might be particularly useful for initial certification due to the exception to the live teaching situation. Students could also be used as test centers similar to the laboratory schools. This would provide a more controlled situation and fewer variables would enter into the evaluation.

A combination of evaluation in student teaching settings and controlled laboratory situations is also an alternative. This might be built into a system where a recommendation from a preparing institution college or universities in addition to testing in a center would be necessary parts of the process for certification. The variations in the testing approach are too numerous to be included in the discussion.

Epilogue

Each model must be considered carefully in terms of the needs identified if criteria there are other issues to be considered for which none are discussed here. The area of certification levels was not presented in the paper and could be an entire area of discussion with about focusing on the value or of models. Another important question is whether or not to the student outcomes as an indication of teaching competence. Success of a practical nature such as job mobility in terms of management, size, size, demographics, and even the economics are examples of valid issues. The questions raised here were more of a philosophical nature and are pertinent to decision making.

The six role described were identified as being along a continuum. This implies that there are many other models which can be considered, but they most likely will differ from these models in degree rather than basic type. Perhaps a system can be developed with positive elements from several of the models described here. It may also be possible that more than one model can be in operation at a given time, particularly if one accepts the notion that certain areas require or more readily fit into a state testing approach while all other areas fit one of the approved program models. The overriding concern is which model or models best serve the purposes of certification.

Performance-Based Teacher Education: Evaluation Issues*

by

[redacted]

Introduction

In this paper I shall consider two levels of evaluation in performance-based teacher education models. The first concerns the evaluation of the basic concept of performance-based teacher education. The second relates to evaluation issues within the context of an ongoing performance-based approach. I shall argue that the important evaluative issues are not those of selecting appropriate and objective measuring techniques but rather involve the types of judgments which dictate who will be evaluated and on the basis of what performances. Those who are seeking unbiased, practical, "how to" answers from my presentation will be disappointed. My purpose is to argue that "how to" questions are predicated upon a prior set of evaluative decisions. Decisions arbitrarily made, ingenuous and powerful than any set of evidence gathering technologies, no matter how objective or complete these are.

The necessity of evaluative issues raised by performance-based teacher education are not new. What is new and significant about performance-based teacher education is that for the first time, systematic evaluation of teacher training models and products is called for. It is the required synthesis of evaluative problems which accrue as foci from several largely individually adopted studies evaluation on scope in performance-based teacher education.

Evaluation

We must start with a definition of evaluation, since the term has multiple and single according to the context within which it is considered. Perhaps the most commonly understood meaning of the term is "a judgment of the values to which learners have assessed the effectiveness of instruction." While this definition is adequate [illegible] goals reflecting the content of those instructional programs. It is limited in that it makes [illegible] procedure for [illegible] the intended outcomes of instruction. Second, a language of [illegible] information-gathering about the process of an instructional [illegible]. Finally, it makes no [illegible] with respect to [illegible] evaluation. To date evaluation of performance-based programs is [illegible] evaluation [illegible] to establish the basis of the [illegible] goals and competencies inherent in the approach. It is important, particularly in [illegible] stages of conception and implementation, to ascertain [illegible] the evaluative issues inherent in performance-based programs. To demonstrate this, take evaluation as used to [illegible] or accompanying a more general paradigm: [illegible] [illegible] evaluation [illegible] [illegible] of [illegible] outcomes. In the context of this paper, these [illegible] [illegible] [illegible] to value judgments of worth or [illegible]. Clearly, this is a more [illegible] definition than the first discussed above. It is [illegible] to [illegible] [illegible] such as the [illegible] [illegible] [illegible] [illegible], or [illegible] [illegible] [illegible]

[illegible] to performance-based teacher education, and uses a perspective to select and evaluate the success of a group of local projects for effectiveness, economy, and component relevance. It assesses the 50 percent of the local program for comprehensiveness in developing the teacher. It is also called upon to judge the quality of products of teacher that are also specific value to participants. The [illegible] to evaluate a different dimension and emphasizes specifying those state evaluation planning process and the performances.

Projects can evolve in the evaluation, determine whether a given evaluation, worthy, and desirable are made use of in the research process, the entire group from the individual philosophies and bodies of literature. It is now to call one of the three examples of school programs which are based, in combination evaluated findings. It may, on the one model to look one of a school. Emphasis on the and against performance-based education could, and will continue to evolve as the other process. The questions asked in the studies are one how performance is differentiate accountability terms but whether we should have them in any reality. [illegible] whether teachers can be evaluated on the basis of their students' performance, and whether they ought to be. There are major efforts [illegible] to differentiate and make distinctions, and it is in these relationships that evaluation comes in as a management. It is important to consider evaluation issues as they relate to implementation and against the basic ideas of performance-based teacher education.

Evaluation of the Performance-Based Model

The questions of performance-based teacher education are many, but they all revolve around a single issue: accountability. There is a concern among some in the success and ability of American schools to support education. Educators are using them as an [illegible] and their activity and products. Evaluation of teachers, and is one aspect of that. Planning [illegible] is becoming a prerequisite for continuing financial support. We want to differentiate and accountable. It will be agree. The assessment tasks for skilled, experienced, and it is how it what educators accomplish are the requirements.

Now, we experience with the that there are the requirements of legislators and administrators, and of accurate teachers. Accountability is a reality. Many forecasting of these who are in the [illegible] approaches. There is growing evidence to the [illegible] state Board and Morgan, 1973), that performance teachers could have value in accountability in the state involved [illegible] of what is their predominant support are. Moreover, performance-based process accountability, [illegible], or accurate, powerful, or a method of meeting the legal of State or establishment unless there is a shortage of assistance and difficult to define [illegible].

Performance-based teacher education is here today as the result of a groups-such as support from teachers, the business administrators, and politicians. It is a more rational, efficient, and accountable method of training and certifying teachers, as well a less so. Performance-based programs are with as the cases they are believed to be better to those. But there are, as in the studies, are because [illegible] problems in methodological methods the opportunity. We have seen that participants favor the way demonstrate its value. The [illegible] to a performance-based approach to [illegible] with its value of legislators and administrators, the student evaluation has therefore, already taken place. The evaluations are not found upon itself, does that rather serve a comparison of the operational values inherent in performance-based teacher. What appropriate in teacher education. As a addition of some of [illegible] does the performance-based approach has to be [illegible] to address.

The broader, proposal for a performance-based approach does not apply specifications of what conditions for use are cases, [illegible] and [illegible] assessments, accountability, are and accountability of a the implementation of this evaluation, which difficult to specifications for special programs. We perhaps more appropriately, it is the value [illegible] apply these measures should of [illegible] the performance [illegible] in assessing the performance [illegible] with and in [illegible] a series of an evaluation design in setting teacher training, and teacher evaluations. It is important to include these accomplishment because, when performance-based teacher education is [illegible] and [illegible], it will be primarily on the basis of the different accomplishment.

Before addressing these considerations, it is appropriate to assess how it is established for the main goals of the proponents that say, in the social sciences, experience that will question [illegible] and practice as [illegible] problem. It gives the [illegible] values [illegible] to the [illegible] and practices. Because the social sciences have been in use, paradigms or models which do not [illegible] accurate, and [illegible] may [illegible] as in [illegible] in this cases or training goals. The [illegible] sciences are different from the social sciences [illegible] in the physical sciences [illegible] have, theories, and [illegible] is a result of physical [illegible] [illegible] that [illegible]

[illegible] [illegible] [illegible] [

Campbell, D. T., Reforms as experiments, American Psychologist, 24, 4, 1969, pp. 409-429.

Combs, A. W., Some basic concepts for teacher education, The Journal of Teacher Education, 23, 3, 1972, pp. 286-290.

Doll, W. E., A methodology of experience: an alternative to behavioral objectives. Paper delivered at the annual meeting of the American Educational Research Association, New York, N. Y., February 1971.

Eisner, E. W., Educational objectives: help or hindrance. Paper delivered at the annual meeting of the American Educational Research Association, Chicago, Ill., February 1966.

Good, T. L. and Brophy, J. E., Looking in classrooms, New York: Harper and Row, 1973.

Jackson, P. W., Life in classrooms, New York: Holt, Rinehart & Winston, 1968.

Popham, W. J., Probing the validity of arguments against behavioral goals. Symposium presentation at the annual meeting of the American Educational Research Association, Chicago, Ill., February 1968.

Popham, W. J. (ed.), Criterion-referenced measurement, Englewood Cliffs, N.J.: Educational Technology Publications, 1971.

Provus, M. M., Evaluation of ongoing programs in the public school system. In R. Tyler (ed.), Educational evaluation: new roles, new means, sixty-eighth yearbook of the National Society for the Study of Education, Part II. Chicago: University of Chicago Press, 1969, pp. 242-283.

Stufflebeam, D. I., *et al.,* Educational Evaluation and decision making, Itaska, Ill.: F. E. Peacock Publishers, 1971.

Tyler, R. W., Basic principles of curriculum and instruction, Chicago: University of Chicago Press, 1950.

Tyler, R. W., Constructing achievement tests, Columbus, Ohio: Ohio State University Press, 1934.

The State Of The Art In Performance Assessment Of Teaching Competence

By

FREDERICK J. McDONALD

To describe and evaluate the state of the art in performance assessment of teaching competence, we should answer three questions: (1) What are the goals of the systems to measure teaching competence? (2) What progress have we made in developing the procedures required for systems measuring teaching competence? (3) What problems must be solved to create effective and efficient systems for assessing teaching competence?

Before attempting to answer these questions, I should first openly acknowledge the assumptions that influence my ideas and those that do not. I assume that even small progress made in assessing teacher competence will be a great improvement over our present evaluations. Because I assume this, I am willing to urge the use of procedures and systems which at the present time are limited or even defective—since I also assume that as we use what instruments and techniques are now available we shall learn more about the nature of teaching competence and progressively improve our methods for evaluating it. I further assume that the cost of measuring teacher competence will not necessarily be prohibitive, despite the fact that many people seem to assume the contrary. It is obvious that instituting a measurement system where there was none before will impose heavy costs. But until we have a complete cost/benefit analysis, I am not willing to assume that these costs will be unmanageable. Further, I do not assume that the evaluation system will necessarily be complex because we do not know as yet what information is needed to make reliable and valid judgments about teaching competence. I choose, therefore, to be optimistic about both the cost and complexity of an evaluation system.

It seems to me that the most reasonable position to take at this time is to acknowledge the many problems and to regard their solution as dependent upon the ordinary processes of research and development. It is self-defeating and destructive to so negatively prejudge the situation that work which might resolve the problems is either impeded or prematurely terminated.

The Goals of an Evaluation System

Let us now return to the three questions with which I began this presentation. The first is obviously the most important: what should be the goals of a performance assessment system? The ultimate goal of such a system is to provide information that enables the decision makers—teacher educators, certification officers, and administrators—to decide whether a teacher candidate has sufficient competence to be permitted to teach. Everything else that we will want to say about this goal is either a refinement of what is meant by making a judgment about competence or a problem that must be solved if such judgments are to be made. I will therefore list a set of conditions that must be satisfied for this goal to be attained.

First, measurement procedures used in the evaluation of teaching competence must have high validity. Validity in this case means we have demonstrated a substantial relationship between a teaching skill or performance and its effects upon students.

At this point in time we know very little about what skills or performances have demonstrable effects on student behavior. The highest priority in a research program on teaching competence must be given to solving this problem. The conclusion, however, should not be drawn that we must defer the development of an evaluation system until all the relevant research has been done. We already have an abundance of ideas on pertinent teaching competencies which we can begin to measure—a necessary first step, and one whose effect on teaching performance can be studied systematically as part of the process of developing evaluation systems.

We cannot or ought not, however, make definitive judgments about teaching competence until the validities of specific teaching skills or performances have been established.

Second, we must establish the reliability of assessment procedures that will be used in evaluation systems. Reliability in this case means the stability of a teaching skill or performance. Does the teaching skill vary considerably from day to day, class to class, topic to topic? We need to know the conditions under which it is most likely to be variable and those under which it is most likely to be constant.

While considerable attention is given to the problem of validity, practically no attention has been given to the performance characteristics of teaching skills. There is some intuitive understanding that a variety of conditions may affect a teacher's skill, but formal research on the problem is nonexistent. It makes no difference how research on this aspect of teaching performance comes out because we need all kinds of information in making judgments about competence. If, for example, a particular teaching skill requires special conditions to elicit it, we must ask whether those conditions were present when the teacher's possession of the skill was being evaluated.

Third, we need to know what, for lack of a better term, I will call the "learnability" of a skill. It is conceivable that a particular skill, even though it has been demonstrated to have highly desirable effects upon children's learning, can be learned by only a few individuals. We do not expect everyone to become theoretical physicists or concert pianists or gold medal-winning athletes because we know that some special aptitudes may be required for achievements of this order. Similarly, there may be forms of teaching whose acquisition demands unusual aptitudes. Thus, we should avoid the trap of assuming that every skill demonstrated to have desirable effects on students will necessarily be required of every person who wishes to teach.

We must develop the kinds of information described in these three points—relating to the reliability, validity, and learnability of teaching skills. We cannot achieve our goals unless we can define what is to be measured, under what conditions it is to be measured, and how the information yielded by the procedure is to be interpreted. But there are some other conditions beyond these three that must be met by any system for measuring teaching competence.

A fourth condition, for example, is that the information gathered about teachers must be as uncontaminated by subjective biases and political processes as is humanly possible. Any observation of human behavior is, of course, influenced to a degree by the characteristics of the observer and by the characteristics of the measurement procedure itself. We shall need to develop new strategies for collecting information on teaching competence if we are to minimize the problems of gathering such information.

Fifth, the conditions of measurement must provide comparable information on groups of teachers. We must standardize the conditions under which teacher behavior is measured. Otherwise, it is impossible to tell whether variations in teacher behavior are due to situational differences or to differences in teacher competence, thus exposing the candidate to the vagaries of idiosyncratic favoritism.

The conclusion should not be drawn, however, that the assessment system should ignore variations in the conditions under which teacher performance occurs. As I stated earlier, it is essential that we determine those conditions under which a teaching skill may be expected to vary. Such information is necessary in evaluating teaching competence. But we cannot treat teaching as if it were so different on each separate occasion that we can never evaluate it. The conflict between establishing reasonable expectations for teaching performance and the variety and complexity of the situations in which teaching occurs is one of the most important problems we have to solve. Until it is solved, our decisions about competence must necessarily be tentative.

Sixth, we must develop an evaluation system that makes due allowance for the teachers' opportunities to acquire the relevant teaching skills. A decade may well pass before teacher education programs have become performance-based and before a reasonable amount of performance assessment has taken place. To demand competence of individuals who have not had the opportunitites to acquire that competence is immoral, and to make judgments that affect their careers on such an insubstantial basis is equally immoral. Assessment should *not* be used for evaluative purposes when the individuals being assessed have not had the opportunity to learn those things for which they are being evaluated.

In listing these constraints, I have outlined the major problems that must be solved if a system of assessing teaching competence is to be developed and used in a practical way; indeed the solution of the problems depends on our meeting these conditions.

BEST COPY AVAILABLE

The State of Measurement Technology and Knowledge

Despite popular opinion to the contrary, measurement technology follows the wants of its users. For three-quarters of a century decision makers of one kind or another have wanted to assess what candidates for teaching positions know. Measurement technology for assessing academic knowledge thus became highly developed. Consequently, we now have widely available tests of knowledge of subject matter and of knowledge about teaching methods. But until recent years there has been no demand of any consequence for the development of performance assessment. As a consequence, this aspect of measurement is largely undeveloped.

We have developed observation scales which permit us to report what teachers do under ordinary teaching conditions; but none of these observation schedules or methods meet any of the constraints which I outlined previously. The conclusion should not, however, be drawn that the use of observation schedules should be terminated; nevertheless, they and any other measurement procedure must be subjected to the kinds of research and development necessary to make them useful within the stipulated constraints.

The basic problem of performance assessment technology is to create a controlled situation or to use an ordinary teaching situation with sufficient knowledge of its characteristics so that their influence on the performance can be assessed. Observation schedules fall into this second group of procedures. The problem of assessing the influences in a wide variety of teaching conditions and performances is enormous. I have come to the conclusion that a more profitable line of research should use controls, simulated teaching situations.

Rather than describe the adequacy of available measurement technology, a rather depressing picture, I shall describe the work in progress on the assessment of performance using simulated teaching conditions. This work is relatively new in the field of teaching competence, and it is conceivable that it may never develop into a practicable technology. But some of the problems must be solved and are more likely to be solved efficiently and satisfactorily by using simulated teaching conditions. The relationships between teaching performance and student learning, for example, can be studied very efficiently by using microteaching sessions. Similarly, the reliability of teaching performances under controlled conditions can also be easily assessed.

We are attempting to build a system that provides

Frederick J. McDonald

information about basic teaching skills, teaching performances, and teaching strategies. This system is designed to use simulated teaching situations of varying degrees of complexity. Initially, both the teaching skill measured and the conditions under which it is measured are relatively simple. The prospective teacher moves through the system from this simple initial level to one where he or she must use a variety of skills, performances, and strategies in combination and must enact them over a period of time. The critical feature of the component systems is that the conditions are carefully controlled. In what ways are they controlled? For any given component, the teaching objectives and teaching performances are specified; for example, if we are attempting to assess the teacher's ability to reinforce students participating in a discussion, the objective of the lesson is set as increasing the amount of student participation in the class. The trainee is told that he will be evaluated on the degree to which students participate in a class discussion. Initially, we want to measure only his use of reinforcement procedures to facilitate participation in classroom discussion. In other components and at other times, the teaching situation will be constructed to evaluate the trainee's ability to present a stimulating problem for discussion or to ask questions which elicit comments from students or facilitate discussion by highlighting differences of opinion.

In order to make assessments at this time, certain factors have to be controlled. For example, if we are to

specific the teacher's use of reinforcement techniques, we do not want the teacher preoccupied with the preparation or development of the topic or with the choice of the topic, nor do we want a group of teachers to have different objectives. Therefore, as I noted, we stipulate the objective, and we also provide the teaching materials for the lesson. Further, we control variability of the students by reassigning the students to a teaching assignment and by providing the preparation period for familiarizing themselves with the material to be discussed. It should be noted, and I wish to emphasize, that the factors which are controlled in one type of situation may be systematically varied in another in order to assess the degree of the teacher's skill.

What can be learned by using this technology? These microteaching sessions can be used to assess a teacher's skill under a variety of different teaching conditions. We have tried several formats of microteaching lessons to determine what can be learned by having teachers teach different lessons and different groups of students. In one format, for example, we asked the teacher to teach four different topics to the same group of students; in another, a teacher teaches different students the same topic; and in yet another, the teacher will teach different topics to the same students and then the same topics to another set of students. In these ways we can test how the teacher's skills change under the different teaching conditions. We have already found that inexperienced teachers' behavior varies from topic to topic, to some extent, even though the topics appear to be highly similar in nature and would seem to offer similar opportunities for using teaching skills.

Eventually, we would expect to have a series of representative teaching situations in which it had been shown that certain kinds of teaching skills are highly likely to be useful. An assessment battery would consist of a series of microteaching sessions covering these teaching situations.

Because the microteaching format has several limitations, such as relatively short lessons and small numbers of students used, we developed a mini-course format to use for more complex teaching situations. A mini-course in this context is a short course lasting 1 or 2 weeks. The teachers teach this course an hour a day for one of these periods of time. For these courses we have also developed a complete statement of objectives, course materials, and achievement tests to be taken by the students at the end of the course. Almost 300 preservice teachers have taught these courses so that we have gathered considerable data on teaching performance and experience with the technology itself.

The results of the analyses of these teaching performances indicate that the microteaching performances are relatively poor predictors of the teaching performances in the mini-courses. But, for several reasons, I am not willing to offer this conclusion as wholly substantiated. First, the subject-matter of the short courses is more complex, and segments of any one lesson would be similar to the teaching situations created by the microteaching session. But their combination presents a different kind of teaching problem. Also, in the situation in which we have gathered data, relatively little performance training took place between the initial microteaching and the subsequent teaching in the short courses. Systematic training using the assessment gathered in the microteaching might influence substantially the relationships between the two kinds of performances; that is, where the teacher has sufficient skills, as revealed in the microteaching performance, the instructors would reinforce these skills so that they were more likely to be used by trainees in the more complex teaching. There is still reason to believe that we have created two sufficiently distinctive types of performance situations which will yield equally valuable but different kinds of information. My own opinion is that the microteaching is more useful for assessing the degree to which a teacher has basic skills; the mini-course is most useful in assessing how teachers integrate these skills into complex teaching performances.

As I said earlier, our goal is to build a system that will assess the simple components of teaching competence and also its more complex aspects. We visualize the coordinated assessment system as made up of periods of microteaching used primarily for assessment purposes and short courses or units that can later be embedded in internship and student-teaching experiences. These components would be arranged so that the training institution could be given information continuously about the level of competence being reached by the trainee.

The student teaching and internship experiences can be used to assess daily performance under uncontrolled conditions. They are useful for providing information on what teachers are likely to do in contrast to what they are able to do. The simulation-type teaching situations that I have described should be thought of as teachers assessing what teachers are capable of doing, whereas on-the-job performance tells us what he or she regularly does. We can also use on-the-job observation to assess such factors as teaching style.

I am convinced, after a year's experience with al-

monitoring systems such as these and another year's analysis of the data gathered, that it is entirely feasible to build a useful performance assessment system that can be embedded in any teacher-training program. As such systems are developed, state departments of education can learn what criteria ought to be applied to the assessment of teaching competence. The prototype that we are developing in our research program can be used to generate concepts about the appropriate means to evaluate each training institution's procedures for assessing the competence of their students.

It is obvious to me that the type of techniques being developed can be used with beginning teachers prior to the granting of tenure, and for periodic reassessment of teaching skill. We should not, however, use these procedures at the present time for continuing certification schemes, any more than we ought to use the current observational procedures and administrator evaluations for that purpose; but there is good reason to believe that, in time, assessment strategies can be developed which will be useful both in preservice and inservice assessment of teaching competence.

One question asked frequently is, "How expensive will performance assessment be?" As I observed earlier, there is no adequate way not to answer this question because we do not know what information we shall need to make judgments about teaching competence. My experience indicates that we can gather large amounts of data about teaching performance relatively inexpensively; for example, three 20-minute microteaching sessions yield an hour of teaching performances. A videotape of these performances can be observed to assess many different skills. In fact, the short cuts yields almost more information than can be quickly processed. We have resorted to audiotaping and videotaping about a half hour of each hour session for the very practical reason that this half hour covers three-fourths of the teaching that occurs. I am not convinced that even this much information is needed. We are studying this problem by taking different samples of the videotape performances that we have and finding the intercorrelations among teaching performances across the samples. This technical work will eventually help us find the optimum amount of observation time that we must use in order to make reliable judgments about teaching skills.

While many people are concerned about the cost of performance testing, few discuss the economies that may be achieved by using them. It is very apparent that, when one looks at videotapes of teaching performances, a diagnosis of basic teaching skills can be made very easily, and this information can be used to speed the learning of these teaching skills. We already have sufficient information on feedback and modeling techniques to know that the information generated in a performance assessment system can be incorporated into feedback and modeling systems to increase the effectiveness of training. Thus, we may expect the institution to benefit by achieving economies in training time and the trainee to benefit by achieving competence sooner when he enters upon actual teaching.

Interactive Simulation Devices

I have been describing performance assessment procedures which require the teacher to interact with students in a simulation of actual teaching. There are some other procedures that appear to be promising in assessing potential teaching competence. We have developed a filmed simulation test that portrays a teacher conducting a class. The film is stopped periodically and the viewer is asked to say what he or she would do in the situation. At other parts of the test the viewer is asked to explain what is occurring in the class, and, at some places, he is asked what advice or suggestions he would give to the teacher. We do not yet have data on the predictive validity of this measurement, but it obviously taps the prospective teacher's attitudes toward students, his perceptions of the problems of teaching, and his understanding of some phases of teaching strategies. The test situation is sufficiently lifelike to create great interest on the part of those who view the film, and invariably the film produces some intense discussions about what the teaching portrays, thus suggesting to us that the film also has great potential for training purposes.

Another type of simulation device has been created to depict the kinds of strategies teachers use. This procedure is a gamelike device in which an experimenter plays the role of student in a highly controlled way. The teacher arranges the subject matter in the form of presentations or questions, and the experimenter responds whenever the teacher asks a question. This gamelike situation does discriminate sharply between deductive and inductive teaching styles. It seems likely that the role of the experimenter can be computerized so that the game can be used by large numbers of trainees in a more efficient manner.

Both of these techniques are illustrative of the kinds that are likely to be developed to assess teacher competence. Obviously, their validity in predicting teaching performance has to be established. But the work to

[illegible] suggests that they are probably of sufficient value to [illegible] such as assessing what a teacher is likely to do in a classroom.

The Problems To Be Solved

The work that I have described here is meant to be illustrative of the possibilities of assessing teaching competence. It emphasized that the problems are primarily technical and partially theoretical. We have not yet developed an adequate technology of performance assessment, but the reason for this is related more to the lack of experience with these technologies than to a lack of knowledge of what to do.

The theoretical problems are more complex and will require extensive research and development. The kinds of procedures that I have been talking about here can be advanced to experiments. For example, assessing the relationship of teaching skills to pupil learning. Carefully controlled experiments can be designed in the approximating format to study hypotheses about these relationships. The assessment data can be used to judge hypotheses about the relationship of teaching competence to pupil learning. But until the basic questions about what competencies make a difference have been resolved, the design of assessment technology will be incomplete.

Each of the conditions governing an assessment system, that I outlined previously, presents a technical problem to be solved. These problems are highly interactive in research, and there is no reason to believe that they cannot be solved in a relatively short period of time.

A basic problem is public of an character. Many individuals believe that competencies cannot be measured because they do not want competencies to be measured. The solution to this problem is obviously complex. Certainly the development of efficient and systematic [illegible] assessment systems will weaken the arguments against assessment itself. There will probably always be those who will say that the most important aspects of teaching are unmeasurable. But I, for one, have never been able to understand how this concept can be [illegible] before anybody has tried seriously to measure the more elusive aspects of teaching.

A New Decision

What should state departments of education, who wish to move to performance-based certification, do about the assessment problem? Obviously, an even more relevant [illegible] of research and development work is [illegible] to be in place and even though we are optimistic about the [illegible] in which the research and development can proceed, it is unlikely that [illegible] assessment procedures [illegible] performance assessment will [illegible]. But for those who are convinced that the competencies that will be the criterion [illegible] assessment is more or less the direction that they should [illegible].

I have [illegible] that many assessment problems are language problems and particularly [illegible]. I suggest that rather than think about assessment as a way to [illegible] competencies, [illegible] be assessed, we may [illegible] assessment think about the criteria that we [illegible] assistance systems from which judgments [illegible] about competencies are made. I am quite sure that we [illegible] to [illegible] a good body of [illegible] research established on the relationship of teaching [illegible] competencies to pupil learning. But I believe that we [illegible] that the most practical course [illegible] to set standards for [illegible] the evidence about competencies and then to [illegible] be demonstrated. Assessing these competencies may be difficult [illegible].

I urge on your consideration [illegible] approach because I believe [illegible] task is an [illegible] assessment [illegible] with the assessment problem. The [illegible] of competencies, and the [illegible] of [illegible] assessment [illegible] training and become [illegible] exercise of the faculty [illegible] with more [illegible] to [illegible] specific goals of [illegible]. My [illegible] is also [illegible] to be that [illegible] assessment [illegible] procedures ought [illegible] to be [illegible] a [illegible] that the biases are [illegible] to be [illegible] judgments about competencies are made. If an [illegible] is involved in these decision-making processes we [illegible] the [illegible] to [illegible] make decisions about competencies, [illegible] of how [illegible] guaranteed that the [illegible] are not being made [illegible] in that a methodology [illegible] of factor-[illegible] competencies [illegible] have methodologies [illegible] to [illegible] the [illegible] of a teacher's [illegible] but we have [illegible] approaches [illegible]. The [illegible] for the [illegible] could [illegible] be [illegible] as a [illegible] addressed to the training institutions. "Given [illegible] conceptions of what constitutes competence, whatever evidence have you gathered that [illegible] whether teachers have acquired these competencies?"

As an [illegible], I want [illegible] would adopt criteria [illegible] the evidence to be presented to support teacher [illegible] competencies, the performance-based [illegible] would [illegible] forward quickly. Many [illegible] are under the [illegible] of [illegible] [illegible] are [illegible] programs which have a [illegible] [illegible] have [illegible] [illegible] evidence that the [illegible] [illegible] go through [illegible] programs are in fact [illegible] teachers. This approach would [illegible] the [illegible] [illegible] on the [illegible] [illegible] the [illegible]

PRIE: Proceed With Caution

by

[redacted]

[redacted]

Current Knowledge on Teaching Competencies

[redacted]

Will Research Take Place in PBIS?

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Alternative Teacher Assessment Strategies*

By

W. JAMES POPHAM

At a time when the Nation's entire educational enterprise is being subjected to increasing public scrutiny, it is not surprising that the classroom teacher—generally conceded to be the pivotal figure in most instructional settings—is being evaluated more frequently and more rigorously than ever before. Perhaps this increased stress on teacher evaluation stems from the widespread concern about accountability in education. Possibly it is the predictable concomitant of a glutted teacher market where employers can finally be choosy, when in former years their chief concern was to get an instructor, any instructor, to cover every classroom.

Whatever the cause, concerns about teacher evaluation have become far more pronounced in the past few years than at any time during this century, even though educational researchers have been continually carrying out teacher effectiveness investigations for well over 70 years. The difference in the focus of these activities provides the key to our understanding of why today's typical teacher starts to perspire a bit when someone mentions teacher competence assessment. In the old days, most teacher effectiveness researchers were searching for a suitable criterion variable which, if located, would permit them to isolate independent variables (such as teachers' personality traits, educational experience, or instructional styles) that would contribute to more effective instruction. Such investigations were accurately perceived by teachers as *research* inquiries and, as such, were not viewed as particularly threatening. Even in those instances where the attention of the investigator was clearly directed toward teacher evaluation, few teachers were very concerned. After all, even if defensible assessment techniques were discovered, it was generally held that teacher evaluation efforts would be directed toward helping teachers, never firing them. The American public had great confidence in the Nation's public schools; and, although everyone knew there were differences in the abilities of teachers, a tenured teacher's position was next to inviolate. For example, a recent search [1] of California's teacher employment records revealed that during the last 40 years not one California teacher has been dismissed on the grounds of incompetence. It is small wonder that in the past our teachers have not been too threatened by teacher evaluation activities.

But they are threatened now—and with good reason. Dissatisfied legislators in a good many states are beginning to enact penalty-laden laws which call for more stringent teacher accountability. The most celebrated of these recent teacher accountability laws is the so-called Stull Act (named after its author, Assemblyman John Stull) passed by the California legislature during the 1971 legislative session (Assembly Bill 293). The Stull Act has generated an immense amount of educational activity among California school people,[2] for its implications are serious indeed. The new law calls for the annual evaluation of all probationary teachers and the biennial evaluation of all nonprobationary teachers. The evaluations must be made on the basis, as stipulated by law, of pupil progress according to district-explicated standards of achievement in all areas of study. What has happened in California as a result of the Stull Act is that an attempt has been made to *operationalize incompetence* so that even tenured teachers can be dismissed if they are evaluated adversely. We can expect to see other state legislatures enacting comparable teacher evaluation laws in the next few years, particularly if the California experiment seems to be working.

But even if no more states establish teacher appraisal systems, there is still a strong likelihood that local districts, perhaps buffeted by school board pressures, will set up some sort of teacher evaluation sys-

* An invited working paper for a meeting of the Multi-State Consortium on Performance-Based Teacher Evaluation, New Orleans, February 25-28, 1973.

[1] Personal communication, Research Department, California Teachers Association, Burlingame, Calif.

[2] See W. J. Popham, "California's New Precedent-Setting Teacher Evaluation Law," Educational Research, Vol. 2, No. 7, July 1972, pp. 13-15.

tem. In view of these developments at the state and local levels it does not require much prescience to be able to forecast an increasing need for the technical devices and procedures required for effective teacher appraisal systems.

Although it is generally assumed by most laymen (and many legislators) that educators currently possess adequate devices for use in evaluating a teacher's instructional effectiveness, nothing could be farther from the truth. The history of teacher effectiveness research is replete with failure after failure in efforts to devise defensible indicators of how well a teacher teaches. Space limitations preclude an exhaustive analysis[1] of the limitations of previously tried assessment schemes, but each of the chief contenders, that is, ratings, observations, and pupil test performance, have fatal defects.

Ratings

Briefly, the difficulty with ratings of teacher effectiveness (characteristically supplied by administrators, but also obtainable from students, peers, etc.) is that different raters have different notions regarding what it is that constitutes good teaching. The same teacher who is rated high by one individual because of such factors as "flexible interaction with learners" and "personable, informal rapport with class" may be rated low by another individual because of "poor discipline" and "classroom anarchy." We all use our private value matrix in judging whether good teaching has taken place, and when we try to pool these disparate sets of rater expectations, chaos is the characteristic result. How many times, for example, has a classroom teacher been rated negatively by a principal because the teacher was conducting class in a way other than the manner in which the principal recalled his/her lustrous days in the classroom. Yet that same teacher may receive a positive rating by the district office supervisor who has a different idea of how teaching should be carried on. It has been observed that one person's humorist is another person's smart aleck. Similarly, one rater's Mr. Chips is another rater's Mr. Peepers.

Observations

With respect to systematic observations of the teacher's classroom behavior, we encounter an interesting assumption. It runs as follows: if certain process variables can be found to correlate positively with desired outcome variables, then by ascertaining whether those process variables are present, on that basis alone we can make judgments regarding the desired outcomes. For example, if it is discovered that a teacher's provision of practice opportunities for learners generally results in desirable learner attainment, then proponents of observational teacher evaluation schemes would contend that we can, at least in part, evaluate teachers on the basis of the degree to which they provide practice opportunities.

The trouble with the logic of this approach is its tendency to force one to the position that the process variables scrutinized by classroom observers are not only necessary for securing worthwhile results with learners, but that they are essentially sufficient. For if the phenomena observed; e.g., amount of teacher talk, are viewed as *means* to an end, why not assess the end results directly without encountering the measurement noise associated with the extra assessment step. For although no upstanding classroom observation devotee will ever assert that those behaviors observed are without exception associated with desired outcomes in learners (such as important cognitive or affective changes), the logic of the observation strategy pushes us to place greater reliance on means-end predictive relationships than the current sophistication of our observational techniques permits. If we are really interested in the ends, why not focus our assessment energy on them?

A second difficulty with observation-based approaches to teacher appraisal is that although a teacher may display optimal use of the classroom behaviors called for in the observation system, there may be deleterious factors present, factors not built into the observation structure, whose presence will essentially cancel out the positive features of the teacher's classroom behavior. The only way to head off this assessment difficulty would be to build an observation system so exhaustive that it could pick up all (or most) negative process variables, but by that time the system would be too vast to be practical.

Another difficulty with observational approaches to the assessment of teacher effectiveness is that whereas they might prove useful in identifying some classroom practices which in general will yield beneficial results with learners, the teacher evaluation game demands personal and particular decisions, not general guidelines. A particular teacher working toward particular goals with particular students in a particular setting

[1] For a more detailed examination of the strengths and weaknesses of various teacher effectiveness assessment approaches, see J. D. McNeil and W. J. Popham, "The Assessment of Teacher Competence," Chapter 7, second edition, The Handbook of Research on Teaching, R. M. W. Travers ed., Macmillan, 1973.

may break all the process guidelines and yet achieve superb results. The particularized interaction effects are too subtle for our currently unsophisticated observation systems. There have been several outstanding pro football quarterbacks whose passing form looked abysmal, yet when the receiver arrived at the appointed spot the ball was always there waiting.

Finally, there is considerable danger that when the stakes are high enough (and job security represents a big bet), many teachers will "fake good." Observation evaluation systems are particularly susceptible to such faking, for in these days of openly described criteria we can expect teachers to know what factors will be involved in the observation system. Indeed, any diligent and legally informed teachers organization should be easily able to unearth the observation dimensions involved. Having been apprised of what practices yield positive evaluations, is it so unrealistic to expect that teachers will tend toward the use of those practices when under observation? Of course, if one wished to employ constant monitoring of classroom behavior through such devices as closed circuit television, then such fakeability fears would be vitiated, but by then most schools would have been closed permanently because of the anti-1984 teachers' strikes.

Pupil Test Performance

The chief deficiency with the use of student test performance as an index of teacher proficiency has generally been that the wrong kinds of tests were employed. Since 1900 most teacher effectiveness research in which pupil test performance was employed as a criterion variable involved the use of standardized achievement tests. Since most standardized tests were designed to serve a different purpose; namely, to permit comparisons among individual learners (not among teachers) they invariably resulted in a "no significant different" outcome.

The difficulties with standardized or norm-referenced tests, particularly for teacher evaluation, have been treated elsewhere,[4] but their two most visible defects can be briefly identified. First, since commercially developed standardized tests must serve students throughout an entire nation, the generalized nature of their content coverage is often inconsistent with local curricular emphases. Incongruent measurement and curriculum results in misleading data. Second, certain psychometric properties of norm-referenced tests (such as their heavy reliance on producing among-learner variance) leads to tests which are sometimes insensitive to detecting the results of high quality instruction.

In the past few years the development of criterion-referenced (or mastery) tests offers teacher evaluators an alternative to standardized tests for assessing an instructor's impact on learners. The judicious employment of criterion-referenced tests for teacher evaluation purposes is only beginning to be seriously investigated.

Teaching Performance Tests

In the mid-sixties, the writer had reached a point of frustration regarding teacher effectiveness assessment devices and, after a reappraisal of alternative assessment strategies, had proposed the development of an alternative approach to solving this problem; namely, through the use of a teaching performance test. Two separate projects[5] were supported by the U.S. Office of Education, each designed to develop and attempt to validate teaching performance tests in different subject matter fields. While the rationale underlying the teaching performance test strategy, as well as the detailed results of these two projects are supplied elsewhere,[6] a brief description of the performance test approach can be supplied here.

One of the major difficulties in comparing teachers for purposes of instructor evaluation is that different teachers have different instructional emphases, thereby making across-the-board comparisons misleading. The teaching performance test counteracts this problem by providing an identical task for different instructors; namely, the ability to accomplish prespecified instructional objectives. The teaching performance test is built on the general premise that one chief reason for a teacher's existence in the classroom is to bring about worthwhile changes in students; that is, changes in their knowledge, attitudes, skills, etc. To the extent that this is true, then one criterion by which a teacher should be judged is his or her ability to bring about such changes. By providing identical instructional

[4] See, for example, W. J. Popham, "Domain-Referenced Measurement and Teacher Evaluation," Education Technology, in press: Robert Glaser, "A Criterion-Referenced Test," Criterion-Referenced Measurement: An Introduction, W. J. Popham, ed., Educational Technology Publications, Englewood Cliffs, N.J., 1971, pp. 41-51.

[5] Performance Tests of Instructor Competence for Trade or Technical Education, USOE Cooperative Research Contract No. OE-5-85-051: Development of a Performance Test of Teaching Proficiency, USOE Cooperative Research Contract No. 3200.

[6] W. J. Popham, "Performance Tests of Teaching Proficiency: Rationale, Development, and Validation," American Educational Research Journal, January 1971, 8 (1), pp. 105-117.

objectives for teachers, then giving the teachers an opportunity to accomplish those objectives using whatever instructional techniques they wish, a measure of the teacher's ability to accomplish given objectives can be provided. One might wish to argue that the better achiever of given objectives will also be the better achiever of his/her own objectives, but this is a question which can be answered empirically. If one simply decides that an important criterion of teaching is the ability to accomplish instructional objectives, then teaching performance tests would appear to have some utility in a data-based evaluation matrix.

The steps involved in a teaching performance test are these: (1) the teacher is provided with an explicit instructional objective (and sample test item) along with any background information necessary to become familiar with the subject matter related to that objective; (2) the teacher plans a lesson designed to accomplish the objective; (3) the teacher instructs a group of learners, typically a small group of learners for a short period of time; (4) the learners are posttested with an examination based on the objective. The examination has not previously been seen by the teacher but its nature is readily inferable from the objective (and sample test item) previously given to the teacher.

In the USOE-supported research studies described above, the purpose of developing the performance tests was primarily research-oriented; that is, it was anticipated that these devices would be employed principally for research purposes such as the identification of relevant independent variables. Consistent with that intent, the performance tests involved in these investigations consumed a fairly large amount of learner instructional time, ranging from 4 to 10 hours. At the conclusion of these investigations, it became clear that if teaching performance tests were to prove practical for teacher evaluation or instructional improvement purposes, they would have to be developed for much shorter periods of instructional time. As a consequence, the writer's recent development work with performance tests has featured instruments which take only 15 minutes of instructional time and are designed to be used with small groups of adults or younger learners. These teaching performance tests, frequently referred to as instructional mini-lessons, superficially appear comparable to the microteaching exercises developed at Stanford University some years back. In rationale, however, they are quite different. The Stanford microteaching lessons emphasize the teacher's acquisition of process skills; e.g., good questioning techniques. The instructional mini-lessons referred to here, on the other hand, focus more heavily on the results of the teaching than upon the instructional procedures themselves.

During the past few years teaching performance tests have been employed both in preservice and inservice teacher education settings.[1] Generally speaking, these performance tests have been of the short duration alluded to above; i.e., 15 to 20 minutes in length. But while these devices appear useful in instructional settings, for example, in helping prospective teachers become more facile at accomplishing prespecified instructional objectives, their utility for purposes of teacher evaluation has been largely unstudied.

In a recent paper[2] Glass has proffered the notion that teaching performance tests may have insufficient reliability to permit their effective use in teacher evaluation enterprises. Glass cited several investigations in which the reliability of teaching performance tests was clearly inadequate. Several of the investigations cited, however, had been conducted as doctoral dissertations or by novice researchers. The reliability of teaching performance tests is as yet a seriously unstudied matter. For one thing, the teaching performance tests used in these investigations have been constructed on an almost opportunistic basis, that is, whatever topics, objectives, etc., have come to the investigator's mind. No attempt has been made to carefully delineate the truly critical dimensions in teaching performance tests. Beyond that, only one investigator[3] has carefully attempted to study the reliability of even these ill-defined performance tests. Results of his investigation will be reported by Millman at the 1973 meeting of the American Educational Research Association. Examination of the Millman findings suggests that the reliability evidence, once again, is not encouraging. But, as indicated above, the nature of the performance test employed in that investigation was not rigorously explicated.

When this paper was solicited as one of several dealing with the "state of the art" in the assessment strategies suitable for performance-based teacher education, I had just completed the final draft of an American

[1] W. J. Popham, Applications of Teaching Performance Tests to Inservice and Preservice Teacher Education. A paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 26-March 1, 1973.

[2] Gene V. Glass, Statistical and Measurement Problems in Implementing the Stull Act. Stanford University Invitational Conference on the Stull Act, October 1972, Palo Alto, Calif.

[3] Jason Millman, Psychometric Characteristics of Performance Tests of Teaching Effectiveness. A paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 1973.

Educational Research Association (AERA) paper describing a set of minimal competencies for a performance-based teacher education program.[1] I had even sketched alternative assessment procedures for each of the competencies. Now I just couldn't bring myself to rewrite the paper or even to subtly paraphrase my original paper. I try to restrict my paraphrasing talents to the writing of others, not my own.

Accordingly, in the present effort I have attempted to focus exclusively on the major assessment alternatives for teacher competence appraisal. Since if performance-based teacher education programs cannot demonstrate that their competency-armed products are indeed better teachers, then the performance-based teacher education folk had best fold up their competencies and slip away into the night.

[1] W. J. Popham, *Identification and Assessment of Minimal Competencies for Objectives-Oriented Teacher Education Programs.* A paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 1973.

BEST COPY AVAILABLE

Conference Participants

Applications Of Teaching Performance Tests To Inservice And Preservice Teacher Education*

By

W. JAMES POPHAM

Irrespective of whether the programs are referred to as "performance-based," "competency-based," or "skill-focused," the advocacy of teacher preparation schemes designed to promote measurable capabilities of instructors is beginning to be quite fashionable among teacher educators. A recent survey[1] of California teacher education institutions found that two-thirds of the institutions participating in the survey indicated they were engaged at least partially in competency-based teacher preparation. There is even a multi-state consortium[2] on performance-based teacher education featuring such forward-looking states as Arizona, Florida, Minnesota, New York, Oregon, Texas, Utah, Vermont, and Washington. Surely, if most teacher educators are not caught up in the competency-based game, then we are at least witnessing the pregame warmup.

Whether competency-based teacher education will become something other than one of those ephemeral fads so common in teacher education remains to be seen. Although many teacher educators are quick to join the competency-based movement at the verbal level, few are willing to devote the requisite energy to devising the criterion measures without which the approach is only rhetoric. In the recent survey[1] of California teacher education colleges and universities, no institution reported satisfactory assessment procedures for a competency-based program. As with so many educational innovations, advocacy is less expensive than workable implementation procedures.

* A paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 26-March 2, 1973.

[1] Performance-Based Teacher Education, Vol. 1, No. 1, December 1972, pp. 1, 5.

[2] Multi-State Consortium on Performance-Based Teacher Education, c/o Bureau of Teacher Education, New York State Education Department, Albany, N.Y.

[3] Op. cit.

Teacher Competencies as Enabling Skills

But there is another aspect of the competency-based teacher education movement which is equally intriguing. If you sit in on almost any discussion among proponents of competency-based teacher education, you will hear them describing the teacher competencies they are trying to promote as though they were ends in themselves. As with many recent religious converts, the fervor of these teacher educators for competency identification has become so all-consuming that they lose sight of what the competencies are really supposed to accomplish.

Reducing the problem to its essentials, we can see that whatever competencies a teacher acquires must be viewed as vehicles for making that teacher more effective. More effective, in this instance, means a teacher better able to help learners. Thus, the competencies most frequently identified by performance-prone teacher educators are really *en route* skills which should contribute to the terminal skill of being able to help learners. For example, suppose one of the teacher competencies we are trying to promote involves the teacher's ability to view real or simulated instructional situations and identify the extent to which certain instructional tactics have been employed. We assume that the teacher who can master such a skill will subsequently be able to apply this skill in real instructional situations. Such skills should thus be viewed as precursive to one's becoming an effective teacher.

Now the point of this distinction between en route and terminal competencies is that the bulk of competencies currently viewed as the staples of performance-based teacher education are well removed from those which might legitimately be viewed as terminal. Unless *en route* skills are constantly verified as being actual contributors to terminal skills, then we have little assurance that defensible competencies are being promoted.

Some of the competencies taught by teacher educators should be given to the terminal performances we wish teachers to display. The purpose of this paper is to explore the teacher education applications of one measurement approach designed to assess such a terminal competency. The measurement approach under consideration is the teaching performance test.

Teaching Performance Tests: Description and Rationale

In brief,[1] teaching performance tests work as follows: the instructor is presented with one or more explicit instructional objectives (often a sample test item) and is directed to prepare a short lesson designed to accomplish the objective(s). If the objective deals with a topic presented to the instructor in the teacher, then relevant background information is made available. After planning the lesson, the teacher instructs a group of learners (either children or adults) for a short period of time (e.g., 15 minutes). The number of learners can be as few as a half dozen or as many as an entire class. At the conclusion of the lesson the learners are given a posttest based on the objective. While not previously seen by the teacher, the nature of the test is readily advisable from the objective the teacher has been attempting to achieve. Characteristically, learners are also asked to rate how interesting the lesson was. If such an interest rating is employed, the teacher is apprised of the forthcoming rating and encouraged to design a lesson which not only accomplishes the cognitive objective, but also promotes positive learner interest ratings.

Using these two indicators, an estimate is provided of the teacher's ability to promote prespecified objectives, both cognitive (as reflected by posttest performance) and affective (as reflected by the interest ratings). Now it may be argued that an overgeneralized estimate of instructional prowess is yielded by measuring a teacher's ability to promote learner attainment of prespecified objectives for a small group of learners during a brief time period. True, while perhaps not as representative of the real teaching world as we might wish, there are sufficient parallels with reality that such an assessment procedure may have utility for teacher educators. In particular, since it more closely approximates a criterion teaching skill than does any of the competencies currently being tested by portfolio-assessed or written measures, it may have certain construct validity advantages over competency tests. Up to a number of of the points will (1) to (3) that illustrate applications of teaching performance tests to teacher education programs, (1) provide actual exercises and an assessment procedure by which performance tests have to have applicants, (2) provide some guidelines derived from this experience, and (3) discuss certain problems we should have solved to the use of teaching performance tests.

Applications for Preservice Education

It is fair to say that many of the educational deficiencies which exist in our schools can be attributed, at least in part, to teacher preparation programs that are structured on process. That is, many teachers are taught with conclusions about desirable and existing ways of teaching, without ever verifying what effect these procedures have on children. Many teacher education personnel, in fact, seem to "know" that their efforts really work, in spite of data suggesting that they may be shortchanging not just an approach that we now believe, like the previous one, themselves are addicted for their own sake. For example, open schools are currently in vogue. Two years ago, it was integrated schools. Is it sound policy for such practices to dictate the nature of the teacher education curriculum?

The teacher does in working successfully among with these new or unproven approaches, but rarely does process teachers performing failures. It is not that we want educational research to be run if an attractive instructional process without checking its worth, it is rather we do believe that time after it should be the course we are using to have unproductive procedures.

Hence, as an a method of assessing what appears to be an effective functional process, the implementation process, the program use of teaching performance tests can provide a vis mechanism to have the teacher's attention on the attainment of instructional. Many of a performance test assessment the quality of a teacher's efforts is provided on an issue inherently with learners, both cognitive and affective. It is difficult to discount the effort of the attention we do apply. It is unlikely, if you are a prospective teacher who during a practice a diligent to could a half dozen or more learners (or more learners teaching performance tests are voluntary, which and the first assessment after your lesson is a determination of

[1] D. McNeil and M. J. Popham, "The Assessment of Teacher Competence," Chapter 1, The Handbook of Research on Teaching, R. M. W. Travers, ed., Rand-McNally, in press.

If more than one objective is employed, as often is typically employed, with the earlier topics, a given may be utilized to identify sufficiently more learners.

[illegible] an effective use [illegible] is difficult to say how one should test such [illegible] a tool as [illegible] a course. [illegible] More specifically, it is important to know [illegible] problems on the part of the teacher, students, [illegible] how to use it, [illegible] how to improve the effective [illegible] of instruction as a [illegible].

Application Three
[illegible] Testing the Value of Instructional Tactics

[illegible] brought the foregoing applications [illegible] as a mechanism for [illegible] students' problems on the [illegible] sequence of instruction, they have suggested that [illegible] instructional procedures [illegible] Essentially, such a study [illegible] the case. We can [illegible] would enable [illegible] of the [illegible] instructional processes. The [illegible] is to [illegible] [illegible] techniques [illegible], in such a way [illegible] [illegible] other [illegible] from [illegible] [illegible] [illegible]

[illegible]

[illegible]

[illegible]

Annotated Bibliography
Teaching Performance Test Articles

Bird, T. (1990). The schoolmaster, the researcher and interactive research. *British Journal of Educational Studies, 38*(3), 239–253.

Cochran-Smith, M., & Lytle, S. L. (1990). Research on teaching and teacher research: The issues that divide. *Educational Researcher, 19*(2), 2–11.

Darling-Hammond, L., & Bransford, J. (Eds.). (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco, CA: Jossey-Bass.

Elliott, J. (1991). *Action research for educational change*. Milton Keynes, England: Open University Press.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record, 103*(6), 1013–1055.

Fenstermacher, G. D. (1994). The knower and the known: The nature of knowledge in research on teaching. *Review of Research in Education, 20*, 3–56.

Gore, J. M., & Zeichner, K. M. (1991). Action research and reflective teaching in preservice teacher education: A case study from the United States. *Teaching and Teacher Education, 7*(2), 119–136.

Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Cochran-Smith, M., McDonald, M., & Zeichner, K. (2005). How teachers learn and develop. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world* (pp. 358–389). San Francisco, CA: Jossey-Bass.

Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher, 31*(5), 3–15.

Kemmis, S., & McTaggart, R. (Eds.). (1988). *The action research planner* (3rd ed.). Geelong, Australia: Deakin University Press.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues, 2*(4), 34–46.

Loughran, J. J. (2002). Effective reflective practice: In search of meaning in learning about teaching. *Journal of Teacher Education, 53*(1), 33–43.

Mills, G. E. (2003). *Action research: A guide for the teacher researcher* (2nd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

Noffke, S. E. (1997). Professional, personal, and political dimensions of action research. *Review of Research in Education, 22*, 305–343.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York, NY: Basic Books.

Stringer, E. T. (2007). *Action research* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Zeichner, K. M. (1993). *Educating teachers for cultural diversity*. East Lansing, MI: National Center for Research on Teacher Learning.

effort to study the psychometric properties of teaching performance tests.)

———, "Teacher Effectiveness: New Indicators for an Old Problem," Educational Horizons, Spring 1973. (A readable rationale statement endorsing the use of performance-based measures of instructor skill.)

Popham, W. J., Evaluating Instruction. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1973. (Two chapters of this book are self-instruction programs designed to provide the reader with skills needed to construct and to use teaching performance tests.)

———, "Performance Tests of Teaching Proficiency: Rationale, Development, and Validation," American Educational Research Journal, 1971, 8, 105-117. (This article describes the background of teaching performance tests and the development of such tests in the fields of social studies and vocational education during the mid-sixties. Attempts to validate the tests are reported and the validation strategy analyzed.)

———, "Teaching Performance Tests as a Vehicle for Instructor Self-Evaluation," The National Elementary Principal, Vol. 52, No. 5, February 1973. (This article describes alternative techniques of employing teaching performance tests as instruments for instructor self-appraisal.)

———, "Teaching Skill Under Scrutiny," Phi Delta Kappan, Vol. 52, No. 10, June 1971, pp. 599-602. (This article examines alternative techniques for assessing teacher skill and describes the possible role of teacher performance tests as an alternative to ratings, observations, and pupil performance on standardized tests.)

Appendix

TEAM S MINI-LESSON

CLINIC GUIDELINES

UCLA Teacher Education Laboratory
Fall 1972

Commencing with the second week of the quarter, each student enrolled in team S will be required to participate either as an instructor or a learner in a weekly mini-lesson clinic. The procedural elements of that participation will be outlined below.

One Booklet Required

During the first week of the quarter, when enrollments have settled down to some extent, a mini-lesson clinic assignment list will be distributed. Nine students will be assigned to a particular mini-lesson clinic group. Three of the students in each group will be designated as the *A* group, three as the *B* group, and three as the *C* group. (These arbitrary designations do not reflect the instructors' grading vision and will have no relationship to the final grade earned in the course.) Each student should, without delay, then go to the student bookstore and purchase one copy of the adult form Teaching Improvement Kit, which bears the same letter as the one he or she was assigned. For example, a student in mini-lesson clinic No. 13 who is designated as a member of the B group should purchase a copy of Form B of the Adult Teaching Improvement Kit. Please note that there are some optional Teaching Improvement Kits available in the bookstore for use with children, but that the required Teaching Improvement Kit, one of the three forms available (i.e., *A, B, C*), is an adult kit. With these materials in hand, the mini-lesson clinics will get underway the second week of the course in the assigned rooms.

One mini-lesson will be taught at each clinic, thus the individual assigned to teach on a given date must be present and prepared to teach. Failure to do so will result in a severe grade penalty. Yet, excused absences will occur. Therefore, all students should be ready to teach one week early in case of an unanticipated absence by the regularly assigned teacher.

General Nature of the Participation

To provide a brief overview, an individual student's responsibilities will be described. First, a student will be obliged to instruct the six students in his mini-lesson clinic group who have been assigned other letters than his/her own. For example, a C member of the mini-lesson group would be obliged to teach a lesson from the C Teaching Improvement Kit to the six A and B students in the group. Second, each student must act as a critic for two mini-lessons taught by the members of the mini-lesson group who possess the same letters. To illustrate, a B member of mini-lesson clinic group No. 7 would not only teach one B mini-lesson him/herself, but would serve (along with another B member) as a critic while two other B lessons are being taught to the six A and C students. Finally, each student will be obliged to locate one or more groups of at least three adults to whom each of the three mini-lessons in his or her teaching kit can be taught, using a form of the posttest other than that employed in the mini-lesson clinic. Thus, in summary, each student will teach one mini-lesson to team S classmates, critique two other mini-lessons being taught to team S classmates, teach three mini-lessons to adults other than team S classmates, and serve as a learner for six mini-lessons.

Procedural Specifics

Now, in more detail, here are the step-by-step procedures to be followed by each student. First, consult the mini-lesson assignment sheet to note which group, letter, and which week you have been assigned. Now read the first three chapters of the Teaching Improvement Kit (pp.1-15). You may wish to examine the mini-lesson assignment information in chapter 4. Do not examine the posttests (on green and blue sheets) which are included in the Teaching Improvement Kits. The mini-lesson clinic sessions have not been assigned to influence your grade in the course, but are specifically intended to help you improve your instructional skills. As a consequence, examination of the posttests prior to your teaching the lesson would reduce the likelihood that the mini-lesson would be beneficial to you.

On the day designated on the assignment sheet, you will be obliged to teach one of the three lessons in your improvement kit (1, 2, or 3). On that day you will instruct (for no more than 15 minutes) the six members of your group who have been assigned other letters; e.g., for the B students this would be the A and C students. The two members of your group who have been assigned the same letter will sit somewhat apart from the student group and will attempt to analyze the quality of your teaching plans, activities, etc. At the conclusion of the 15-minute lesson, the perforated copies of posttest 1 (green paper) should be removed from your Teaching Improvement Kits and distributed to your six students. You should not have seen the posttest prior to this moment. These tests should be quickly completed and scored by the mini-lesson clinic group leader. On the basis of the average scores on (1) the interest rating and (2) the posttest, discussion of the teaching should be conducted in terms of the results produced. That is, good learner performance should result in a discussion focused on instructional tactics which seemed effective; poor results should lead to a discussion focused on instructional procedures which might be altered. To assist you in judging how effective a given teaching performance was, at the rear of each Teaching Improvement Kit data are available regarding how successful other teachers have been with each mini-lesson. Results of each student's performance will be turned in, but not for grading purposes, by the mini-lesson clinic leader.

On two occasions, therefore, each student will be serving not as a learner but as a critic. On those two occasions the student (having access to the Teaching Improvement Kit under consideration) will have had an opportunity to examine the mini-lesson assignment prior to the clinic. He or she will undoubtedly have some thoughts regarding an appropriate instructional procedure. This may be beneficial during the post-lesson analysis session. It will be useful to have two people, other than the clinic leader, who are as conversant with the mini-lesson requirements as the teacher. During the analysis discussion, it is anticipated that the mini-lesson clinic leader and the two critics will take primary responsibility for isolating elements of the lesson that were particularly effective or ineffective. Remember that the appraisal of instructional means should be made chiefly in terms of learner results; that is, the averaged interest ratings and posttest scores.

The final responsibility for each team S student is to teach all three of the mini-lessons in his or her kit to another group of adults; that is, someone other than team S students. Since each of the mini-lessons in the Teaching Improvement Kits has two equivalent posttests, the team S students should not, either deliberately or inadvertently, examine posttest 2 (blue sheets). Ideally after the original teaching of that mini-lesson in the mini-lesson clinic, each student should locate a group of at least three adults (friends,

relatives, or people off the street) who would be willing to serve as students for the approximately 20 minutes involved for this assignment (15 minutes of teaching plus 5 or so minutes of testing). The mini-lesson should be taught and results summarized on the Mini-lesson Posttest 2 Summary Report Forms which have been provided (see attachment). The information called for is brief, yet should be completed in its entirety. These Mini-lesson Posttest 2 Report Forms should be turned in weekly during the 9 o'clock class meetings. In other words, an individual team S student should have an opportunity to reteach the mini-lesson he originally taught to team S students after a critique of the first lesson. Hopefully, this analysis will be useful in promoting improved learner performance. The other two lessons in the kit will, for that student, be taught for the first time. Ideally, the discussion of a classmate's teaching of that same lesson will prove useful in helping devise an effective lesson.

Other Considerations

Results of the weekly mini-lesson clinics will be made available so that students who wish to compare their performance with that of other team S members may do so. As indicated previously, there are other, optional Teaching Improvement Kits available in the bookstore for use with younger learners (elementary school age). Furthermore, mini-lessons in the adult kits can be used with mature secondary school students. Thus, it is possible to teach mini-lessons from another kit (other than the letter assigned) to high school age students. Either of these activities; that is, teaching mini-lessons to elementary or secondary school youngsters, should be considered optional for team S students. Remember, the whole purpose of the mini-lesson clinic operation is to improve the team S member's skill in accomplishing prespecified objectives with teaching procedures which are also interesting to the learners. The amount of time that you can devote to promotion of this particular competency will undoubtedly yield great benefits to your future students. As a consequence, please approach the mini-lesson clinics and outside mini-lesson assignments as a real opportunity to increase your instructional skills.

MINI-LESSON POSTTEST 2 REPORT FORM

Your name _____ Mini-lesson _____

Date mini-lesson taught _____ Today's date _____

Setting for teaching and types of learners (one or two sentences)

_____

_____

Number of students taught _____

Results: average interest rating _____

average percent correct on posttest _____

Was this the second time you taught the mini-lesson? Yes _____ No _____

If this was the second time, supply the average interest rating _____

and average percent correct _____ for the first time you taught the mini-lesson.

Comments (optional): _____

_____

_____

Below give the names and phone numbers (if available) of at least three of the students you taught.

_____

_____

Identification And Assessment Of Minimal Competencies For Objectives-Oriented Teacher Education Programs*

By

W. JAMES POPHAM

Classroom teachers have hundreds of things to do. It should follow, then, that prospective classroom teachers have hundreds of things to learn. But too many well meaning teacher educators have used these two premises to draw the conclusion that they therefore have hundreds of things they must teach, and that represents a serious error.

Given the instructional time available in typical teacher education programs, we must be more modest in our aspirations. When we ask teacher education candidates to swing an axe at every tree in the forest, they may fail to fell even a sapling. Far too many teacher education programs are predicated on a cover-the-waterfront concept: that is, give 'em the works in cultural foundations, educational psychology, and instructional methods. And the use of the verb "cover" is quite deliberate. Most teacher educators feel compelled to cover content in their courses that they perceive as pertinent to the teacher's responsibilities. When these professors have covered such content, they sleep easier at night. Few of these coverage-culprits ever verify whether their extensive coverage of subject matter ever results in any payoff for the teacher candidate, other than the ability to pass a memory-oriented final examination. And as in so many content-coverage courses, not just those in a teacher education sequence, what was covered one semester has faded from one student's memory by the first week of the next term.

Particularly at a time when teachers are being weighed more scrupulously on the public's accountability scales, teacher educators who persist in covering all relevent topics are probably doing an injustice to the teachers they are responsible for preparing. It makes more sense for the teacher educator to select a limited number of competencies which teachers should acquire, then focus the program's resources on making certain these skills are acquired.

Now even if this point of view (i.e., a focus on the attainment of a modest set of competencies) were assiduously followed, there would still be considerable disagreement regarding which competencies to promote or, in more general terms, what kinds of content to emphasize. Some would prefer to focus on the teacher's attainment of a wide repertoire of instructional techniques. Others might attend more directly to the teacher's becoming a more integrated human being. Still others would emphasize the teacher's acquisition of subject matter expertise. The alternative emphases are myriad.

An Objectives Orientation

The remainder of this analysis will describe a particular orientation which can be described in general terms as an outcomes-focused approach. An outcomes-focused approach emphasizes the results that a teacher's efforts produce in modifying the behaviors of learners and can be contrasted with more process-focused strategies which attend to the instructional ploys a teacher utilizes with pupils. Because instructional objectives can serve as a convenient way of describing the intended results a teacher wishes to achieve with learners, we may refer to one variant of an outcomes-focused approach to teacher education as objectives-oriented. An objectives-oriented strategy for educating teachers will be treated here.

The rationale for an objectives-oriented approach to teacher education characteristically rests on a central assumption, namely, that the *raison d'être* for a classroom teacher is to bring about worthwhile changes in learners, i.e., important kinds of improvements in their knowledge, attitudes, skills, etc. Proponents of an objectives-oriented teacher education program believe that even if a teacher lectures with consummate skill, but the students are left unchanged, the teacher has failed.

* A paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 26-March 2, 1973.

Similarly, they contend that even if the teacher has had a nondirective discussion with the artistry of Carl Rogers, but the students are basically unaffected, then the teacher has failed. The criticism, quite clearly, is not with the teacher does, but what happens to pupils as a consequence of what the teacher does. Few objectives-oriented teacher education programs are not somehow wedded to this basic view of a teacher's mission.

But how do objectives enter the picture? Well, their chief value is in helping teachers identify more clearly, prior to instruction, the kinds of changes which should be promoted in the learners. Statements of instructional objectives are nothing more than that, convenient descriptors of intended changes in learners. In the early 1960's, advocates of the oft-maligned behavioral objective endorsed such formulations vigorously because of their focus on the learner's post-instruction behavior, not on what the teacher was going to do or the content that the course would cover. It is unfortunate that some educators have become so entangled with behavioral objectives they have made them a fetish. Precise instructional objectives, in the main, are simply statements of what teachers want to happen to learners as a result of instruction. The more explicitly these intentions can be formulated, the better we can tell whether the intentions have been realized, and it is for that reason that most proponents of objectives strive for measurability as the sine qua non of an acceptable objective. But remember the central purpose of an instructional objective—it is to help an instructional planner conceptualize the kind of changes to be promoted in learners.

Proponents of instructional strategies featuring measurable objectives should forthrightly admit that their conception of the instructional process is generally one based on rational decision making. Some critics of an objectives-oriented approach denigrate such strategies as "industrial models" of education and therefore somehow unworthy of man's truly humanistic capabilities. They would prefer less systematic and intellectualized approaches, favoring instead more intuitive, dynamic models. But when Aristotle isolated the essence of man as his rational animality and held that a person's potentials were realized to the extent that these rational powers were actualized, he offered objectives-oriented teacher educators a satisfactory source-argument. To plan one's actions on the basis of the action's likely consequences is less industrial than it is rational. To be clearheaded is not to be mechanistic. To define anticipated outcomes in advance does not relegate one to an assembly line mentality. On the contrary, to be rational in our education decisions will give our students the best chance of prospering from the education we provide them.

Are Instructional Techniques Unimportant?

With most objectives-oriented teacher education programs, it is proper to assert that a distinction is drawn between instructional means and instructional ends, with the stress typically on ends. But as anyone who has attempted to achieve a significant end will agree, it is wrought about by employing appropriate means. Hence, an outcomes-focused teacher educator must be particularly attentive to instructional techniques, enhancing the teacher's skill in employing a wide repertoire of teaching tactics, for it is only through the judicious use of such procedures that significant kinds of results in learners can be attained.

Minimum Competencies

Programed instruction specialists are familiar with an approach to the development of instructional materials known as lean programing. In such a strategy the programer tries to accomplish a given instructional objective with the least possible amount of instructional stimulus material. Aside from its obvious economic advantages, lean programing carries with it a dividend when an early version program is unsuccessful. It is easier to improve a low density program by supplementing it than it is to delete segments of a high density program, for in the latter approach we may be excising the very ingredients that contributed to whatever effectiveness the program possessed.

Similarly, what is being proposed here may be described as lean competency protection, for only those competencies of an objectives-oriented teacher education program will be recommended. Now surely teacher education candidates will learn other things as they complete their preparation programs; they may even learn some of the hundreds of things referred to at the outset of this paper. Since it will be easier to supplement a few minimal competencies than to delete from a more diverse array, it is proposed that only those such skills be emphasized in an objectives-oriented teacher education program.

The remainder of this discussion will isolate these three minimal competencies, offer some support for their importance, and discuss alternative methods of assessing the degree to which each has been attained. Then, these competencies may be used as the guiding goals of either a preservice or inservice teacher educa-

...tion effort. The differences in strategies for promoting the competencies for experienced or beginning teachers are only superficial and the differences in assessment tactics almost nonexistent.

Competency Number One

Since the chief assumption of an objectives-oriented program is that teachers should promote worthwhile changes in learners, it is not surprising that the initial competency to be focused on deals with that basic skill.

1. Teachers must be able to achieve prespecified instructional objectives with diverse kinds of learners.

This competency implies that a skilled teacher should, when presented with clear statements of intended changes in learners, be able to devise instructional sequences which will work, that is, which will bring about the sought-for changes in the learners. Further, the competency implies that this skill be manifest with different kinds of learners, for example, children of differing ages, ability levels, ethnic backgrounds, socioeconomic status, etc.

The truly professional teacher not only will need to be consistent with sound instructional principles in order to design such instructional plans but will have to discover what kinds of teaching tactics personally prove effective. Not all students can get good results from the same fiddle. Different people must adopt different teaching styles. For some teachers a nondirective approach will work beautifully, while for other teachers such a strategy would be a disaster. It is imperative that a teacher discover what communication style, coupled with relevant instructional principles, typically results in the attainment of prespecified instructional objectives for that teacher.

Assessment Tactics

There are two prime contenders for assessing the degree to which this initial competency has been attained. The first of these involves the use of teaching performance tests (or instructional mini-lessons) whose rationale and applications are described elsewhere.[7] Briefly, a teacher is given a measurable instructional objective (typically dealing with a novel topic) along with any necessary background information needed to understand the objective. The teacher plans a short lesson (e.g., 15-20 minutes) designed (1) to accomplish the objective and (2) to be interesting to the learners. The lesson is then taught, either to a small group of six to eight learners or to an entire class. At the conclusion of the lesson, a posttest is administered to the learners. The posttest has not been previously seen by the teacher, but its nature is readily inferable from the clearly stated objective. A form requesting the learner to "rate how interesting the lesson was" is also provided. The teacher is judged on the basis of whether both the cognitive intention (the objective as measured by the posttest) and the affective intention (the promotion of learner interest as measured by the rating forms) have been achieved.

The recency of serious research attention given to teaching performance tests as an evaluative tool probably renders them inappropriate at the present time for assessing individuals other than those at the extremes of a distribution, i.e., the particularly good or particularly poor goal achievers. Performance tests may also be used to evaluate the efficacy of a teacher education program by administering them on a preprogram and postprogram basis to the teachers involved. For example, suppose two teaching performance tests (*A* and *F*) were employed. One of each could be administered to all the teachers (or teacher candidates) at the beginning and at the close of the program. The predictions would be that $\bar{X}_{A_{post}} > \bar{X}_{A_{pre}}$ and $\bar{X}_{F_{post}} > \bar{X}_{F_{pre}}$.

Although the reliability of different teaching performance tests has not yet been established with sufficient precision to warrant their use for individual evaluations, with more systematic delineation of the key elements constituting such tests we may find that in the future they can be used for more fine-grained analyses of individual teacher's mastery of competency number one.

A second approach to the assessment of the initial competency is to allow teachers to post their own instructional objectives, develop a congruent mastery examination, then instruct a group of learners in order to attain the objectives. Interest ratings can also be employed here. Because an objective generated by a teacher can be less readily compared with objectives pursued by other teachers, there is the additional responsibility of the educator to appraise the quality of the teacher's objective, not to mention the consonance of the test with the objective. The advantage of this

[7] W. James Popham, "Performance Tests of Teaching Proficiency: Rationale, Development, and Validation," American Educational Research Journal (in press). W. James Popham, Applications of Teaching Performance Tests in Inservice and Preservice Teacher Education. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 26-March 1, 1973.

[illegible] the first two units mentioned in objectives 3a and 3b of general methodology in a concrete phase objectives.

3. Exercises must be able to show the second-ground effects of their interventions.

In spite of good intentions, most combined with good evaluation-achieving skills, a teacher's efforts will often produce unforeseen consequences and be[illegible]ed with mistakes. Hence, all in the execution of an educational plans be considered. The teacher must be skilled in determining the results of what happened in instances, including in cases when was supposed to happen.

There are several different techniques a teacher might employ to discern unanticipated effects of an intervention, such as the use of an external accompanying participant-observer questionnaires (e.g. "Can be best and worst things that happened to you because of the course"), the unobtrusive observation questionnaires which attempt to collect the opinions and suggestions of those expected to suffer an effectives, which might avoid the[illegible] of interaction, (c) post-participation techniques such as the assignment of an essay to the class dealing with topics such as "My experience in Biology 1" or "Autobiography of a U.S. History Student," and (d) the investigations of the results of a teacher's efforts by a colleague who follows the[illegible] from those in other[illegible] This is why attempts to discover unanticipated consequences of[illegible] when the teacher's objectives were what happened to the students.

It is important that objectives-oriented teacher-education programs possess the third competency, the ability to move to the next stage that teachers and their instructors themselves accomplish their own skills. But at the same time, personal humility and efficacy which stems from control over the[illegible] of tasks.

Assessment Tactics

It is now clear that the first two competencies, the ability to identify important to[illegible] and to state them in[illegible] behavioral dispositions, is central to the assessment phase of instruction.

The[illegible] preferred method of getting at the dispositions is to employ an instrument such as that presented in the appendix. The relevant part reading[illegible] of this instrument, "Learning to Learning," is explained[illegible] and the[illegible] he[illegible] is asked to respond various degrees of agreement with a series of statements regarding instruction, some of which deal with the use of[illegible]-and-pencil test[illegible].

But the real test of the competency, the ability of[illegible] a teacher to describe, in terms easy to grasp by[illegible], a teacher who[illegible] to use[illegible] such effects.

Perhaps[illegible] approaches offer the greatest promise with respect to ascertaining whether this competency has been attained. Instructional situations could be presented to the teacher either on paper or[illegible], in film, in which there are clearly stipulated objectives plus some evidence as to the degree to which they had been achieved. In addition, there would be some subtle[illegible] unanticipated effects of instruction. The teacher would be asked to[illegible] the worth of the instruction and a second would be asked of the extent to which attention to the unanticipated side effects had been incorporated in the instruction.

Existing teachers to describe their general evaluation strategies in[illegible] circumstances. For[illegible] could then be given both descriptions to see if, in essence, in this highly[illegible] situation, unanticipated side effects were built into the teacher's evaluation plan.

A Beginning

In[illegible], an attempt was made to this[illegible] to[illegible] the preparation that many competencies should be built[illegible] the[illegible] exercises for teacher education programs. The objectives-oriented teacher education and both was described and then developed in[illegible] context for such a strategy were outlined[illegible] with alternative[illegible] to such[illegible].

Three basic competencies were identified on the basis of the writer's experiences with[illegible]-focused teacher education programs. They are attributed on the belief that teachers who possess such skills will be able to do a better job for the learners they[illegible] to serve.

The assessment tactics, however, are presented not as[illegible] to be[illegible] tools. Regrettably, the development of[illegible] assessment[illegible] that[illegible] what objectives-oriented teacher education is about[illegible] not necessarily[illegible]. More importantly, perhaps, it may encourage teacher educators, both objectives-oriented and those of other persuasions, to[illegible] the[illegible] of[illegible] with[illegible] in their programs and the[illegible] which they employ to[illegible] them[illegible].

Appendix

Looking at Teaching

[illegible]

This document consists of brief descriptions of instructional situations, each of which is followed by two questions. Please respond to each of the questions by circling the appropriate letters on the list of possible responses according to the following schedule:

SA — Strongly Agree
A — Agree
D — Disagree
SD — Strongly Disagree

Please use the right to circling answers in this instrument. In responding, do your best to answer each question in relation to your view of instruction. Remember, please try to answer as accurately as your impressions.

[illegible]

Mr. Bell is a senior high school history teacher who believes very strongly in "open education." He arranges desk in groups of five, has numerous environmental cues in his room, emphasis on discussions, little individual reports of answers, projects students have developed on walls of their personal interests. Mr. Bell feels that students are generally cooperative in his classroom, but some of these cooperative accommodations that they do not like learning enough to pursue. Some few consider high school history boring.

SA A D SD	1. Mr. Bell has no right to implement open education if it requires students to accomplish course coverage.
SA A D SD	2. The use of techniques like Mr. Bell is practicing will generally be welcomed by its students.
SA A D SD	3. It is impossible to combine the best of open education with assigned course coverage.
SA A D SD	4. Mr. Bell should have no truck on open environment plans, unless the he was hired, prior to the implementation of the measure.
SA A D SD	5. Mr. Bell should not be in direct any effects of the instructional choices other than those he personal might encourage.

[illegible]

An experienced school teacher, Mrs. Price, recently made some moves in fourth-grade children. Accordingly, many students who came in her class and took gains with her 30-37 percent pupils. Mrs. Price's difficult self-contained beginning students for three years children were encouraged them to go to the school develop the atmosphere that was in the Mrs. Price was students their learning abilities. Although the performances of these children is noticeable, they have become independent thinkers, they are adequate in communicating ideas, by the good readers in the class interested in their participation in the learning centers.

SA A D SD	1. Self-contained classrooms are the most valuable structure for any teacher.
SA A D SD	2. Mrs. Price should have more confidence in the good student's ability in her class because their performances were the responsibilities of previous teachers.
SA A D SD	3. Mrs. Price thought it was not fortunate that Mrs. Price should realize that the long term effects of the stress Mrs. Price's may have been more harmful to the good students than those in teaching.
SA A D SD	4. It is unclear for 30-37 percent of children in most fields, in any grade like Mrs. Price can get left to do the most valuable.
SA A D SD	5. Mrs. Johnson says Mrs. Price's experiences a certain amount of unnecessary those content and high proficient.

[illegible]

Mr. [illegible] is a high school English teacher who

[illegible] [illegible] [illegible] [

in ongoing school settings, and that it meet designated performance standards.

While such meanings may sound complex and strange, they are not at all arbitrary, for they follow necessarily from the basic definition of competence as the ability to bring about the outcomes expected of a certificated job position. This is so because jobs are situation specific. If such meanings are hard to understand the reader is referred again to the technical paper that spells out in considerable detail the implication of a competency definition of this kind for program operation (Schalock, 1973a).

The OCE Definition of Gradualism in the Demonstration of Teaching Competency

As indicated previously, the "Oregon Process Standards" call for competency demonstration at the level of initial certification to take place in a full-responsibility teaching situation of no less than 2 weeks duration. The assumption underlying this expectation is that a demonstration context of any less demanding nature would be insufficient as a source of evidence about ability to bring about the expected outcomes of teaching. A year of full-responsibility teaching is suggested as the demonstration context for competency assessment at the level of basic certification and 2 to 3 years of full-responsibility teaching, after basic certification has been achieved, is suggested as the demonstration context for competency assessment for the standard certificate.

The successful performance of teaching functions under conditions of full-responsibility for a 2-week period is a demanding task. It is also one that is not likely to be performed successfully without prior experience in teaching. The experimental program at OCE recognized this fact, and was planned to incorporate an extensive system of precertification teaching in which teaching competency could be demonstrated within a series of graduated competency demonstration contexts. Two demonstration contexts were to be provided prior to the 2-week context in which competence was to be demonstrated for purposes of initial certification, lesson teaching, and short-term (2-5 days) full-responsibility teaching. The lesson teaching context was to require the preparation and presentation of three lessons on three separate days and in three different subject areas.

Both demonstration contexts were to be provided within what amounts to the first two terms of a three-term professional preparation sequence. Observation and informal lesson teaching were to precede the preparation and presentation of the three lessons that were to receive formal assessment. Performance standards for lesson teaching were to be based upon a summary of performance in the three lessons, and were to be met before a student was to engage in short-term, full-responsibility teaching. The ability to bring about desired learning outcomes in pupils was not to be required as a competency at the level of lesson teaching, though the assessment and display of the learning outcomes achieved in lessons were.

After competence was demonstrated in lesson teaching, a student was then to demonstrate his competence in short-term, full-responsibility teaching. Essentially the same competencies demonstrated in lesson teaching were to be demonstrated in the new context, but with more demanding performance standards. Some additional competencies were also to be demonstrated. Performance standards were to assume much the same form as they did in lesson teaching, including not being held accountable for bringing about desired learning outcomes. That was to be a requirement, however, for competency demonstration at all levels of certification.

The demonstration of competence at the level of short-term, full-responsibility teaching was to permit a student to move on to student teaching, to intern teaching, or to a student teaching equivalency examination. The basic assumption underlying the design of the graduated demonstration contexts to be included in the program was that if students were to perform effectively in a 2- to 5-week, full-responsibility teaching situation (student teaching) that provides the context for competency demonstration at the level of initial certification they would have to have a carefully tailored set of precertification teaching experiences that would prepare them to do so.

The Utility of Competency Measures For Instruction and Job Placement

From its inception the assessment system at OCE was seen as serving the purposes of instruction, certification, and job placement. Accordingly, in each step in its development the system was to be influenced by the needs of instructional staff, the certifying agency, and the personnel officers of school districts. Information obtained through the assessment system was to serve online instructional (supervision) decisions as well as certification and hiring decisions. The latter two sets of decisions were to be facilitated by the preparation of "competency profiles" that depict competencies demon-

strated in each demonstration context. To make the preparation of such profiles feasible they were to be prepared by computer.

The Utility of Competency Measures For Research

One of the major problems that has confronted research on teacher education, or for that matter research on the effectiveness of instruction, has been the lack of adequate measures on the outcome side of teaching. One of the major aims for the OCE assessment system was the development of a system of measurement that would eliminate this problem, and by so doing enable a first-rate program of research to be established at the college. In order for this to happen a long-range, systematically designed program of research was planned on the quality of the measures themselves. The results of the first year of this research, and the lines which it is to take during the coming year, are described in subsequent sections of the paper.

The Feasibility of the Assessment System, Given the Resources Regularly Available to the Teacher Education Program at OCE

While help was to be available for the development of the proposed assessment system through grants from Teacher Corps and the National Center for the Improvement of Educational Systems, U.S. Office of Education, the system was seen as having to operate within the constraints of the resources ordinarily available to the teacher education program at OCE once it was developed. Since the system obviously would require resources to operate, this meant that ways would have to be found to utilize resources presently available in different ways, or to find as yet untapped resources and enlist them in support of system application. This stance was consciously adopted for two reasons. First, OCE did not wish to develop an elaborate assessment system and then find itself in a position of being unable to operate it because it was too costly. Second, if the system was to have general utility to the field it would have to be functional within the resources ordinarily available to most teacher education programs.

The OCE System for Assessing Teaching Competency, Year I

Starting with the view of system requirements that has just been outlined, work was begun on the actual development of the system in September 1972. Work on the system progressed simultaneously with the development of the experimental program as a whole, and as with any interdependent effort, steps taken in one arena both influenced and were influenced by steps taken in the other. Only the results of this developmental process will be described in the present paper. The process itself is being chronicled in a monograph that describes the development of the experimental program at OCE historically (Schalock, Kersh, and Garrison, in preparation).

Teaching Competencies to be Demonstrated

In order to develop a system for assessing teaching competency one has to be clear about the teaching competencies to be assessed. Fortunately, the instructional staff at OCE were reasonably clear as to the competencies that they wished to see demonstrated at the level of initial certification, so this step in the developmental process was accomplished with reasonable ease.

It will be recalled that the definition of competency adopted at OCE, and in Oregon generally, has to do with the ability to bring about the outcomes expected in the performance of a role or function within a certificated teaching position. As used in this definition, role or function refers to the "largest meaningful classification used in describing units of work within a teaching position" (p. 19 of the April 12 draft of the "Oregon Process Standards"). The teaching functions accepted by the OCE staff as a point of departure in developing the assessment system were as follows:

- Defining the objectives of instruction
- Adjusting instruction for the individuals involved (teacher and pupil)
- Selecting appropriate materials and procedures for instruction, given the objectives and individuals involved
- Organizing the learning environment to support instruction
- Interacting with pupils (for pupil success) in the process of instruction
- Evaluating student growth (cognitive and attitudinal)
- Defining next learning steps, and the instructional procedures that attend them, given all of the above.

These seven functions were proposed initially by Dr. Herbert Hite, now chairman of the Department of Education at Western Washington College, as content for

BEST COPY AVAILABLE

the ComField model (Schalock and Hale, 1968). They served as the primary organizers for instruction and assessment in the experimental program at OCE during its first year of operation.[1]

In developing the assessment system around these seven teaching functions (competencies to be demonstrated), the decision was made to obtain evidence as to how well the functions were planned as well as how well they were carried out in actual teaching. This gave rise to two reasonably distinct sets of assessment procedures, one focusing upon the assessment of lesson or curriculum plans and the other upon lesson or curriculum presentations.

H. D. Schalock

Competency Demonstration Contexts

In a program that defines competency in terms of the performance of teaching functions in an ongoing school setting the identification of the contexts in which competencies are to be demonstrated becomes as critical as the identification of the competencies themselves. In fact, it has been the OCE experience that it is not possible to think in terms of competency without thinking in terms of the context in which competency was to be demonstrated! Until these two aspects of competency definition were sorted out, little progress was able to be made in the development of the assessment system.

As indicated previously, the first year of work on the system saw the development of assessment capability in two demonstration contexts: lesson teaching and short-term (2-5 days), full-responsibility teaching. Some work was also done in relation to a student teaching equivalency demonstration context, but this was not completed. The first two of these demonstration contexts deal with precertification competency demonstration, and are designed to prepare students gradually to demonstrate the level of teaching competence required for initial certification. The third is designed as the context in which competence at the level of initial certification may be demonstrated.

Reasonably detailed specifications surround each of the competency demonstration contexts. Those surrounding lesson teaching illustrate the form and substance that they tend to take.

In partial satisfaction of requirements for completing the two-term block sequence that comprises the 1972-73 Experimental Teacher Education program at OCE, each student is to demonstrate competence in teaching elementary pupils in appropriate school settings. Each ETE student will accomplish this objective by assuming full teaching responsibility for two to five successive school days in an ongoing program of instruction, in such fashion that their performance is judged adequate by the student, the college supervisor, and the classroom supervisor in accordance with the standards set for performance in such a situation.

In preparation for the full-responsibility teaching demonstration, each student will prepare and teach three lessons, one in each of three different subject areas. The preparation and presentation of each lesson shall be done under direct supervision and in such fashion that a complete assessment can be made of strengths and weaknesses, and that the student may profit from the assessment. Satisfactory completion of all three lessons is a prerequisite to entry into the full-responsibility demonstration. The three teaching lessons are to meet the following conditions:

1. At least one lesson shall involve the teaching of reading, and at least one the teaching of art, physical education, or music;
2. At least one of the lessons shall enable the student to demonstrate ability to teach children from diverse cultural or ethnic backgrounds;
3. At least one of the lessons shall have as its primary or secondary instructional objective the learning of career awareness;
4. The desired learning outcomes for each lesson should be limited so that a lesson normally will not require more than sixty (60) minutes of instruction and not less than twenty (20);

[1] Teaching functions, of course, are always imbedded in a subject matter context, and as a consequence their assessment requires an accompanying assessment of the adequacy of the content they carry.

5. Each lesson shall be taught to a group of students numbering five or more, and at least one lesson shall involve an entire classroom or its equivalent;
6. The student will be evaluated on the basis of a specified set of teaching tasks or functions (complex skills) within each lesson, ranging from the planning of the lesson through the assessment of pupil learning from the lesson;
7. Formal lessons may be prepared and taught at any time during the two terms allotted the experimental program, though the student needs to remember that if all requirements of the program are to be met within the two-term period, the three teaching lessons will have to be completed in time to permit the student to arrange for and complete the full-responsibility teaching demonstration;
8. An objective summary of a student's performance in the preparation and presentation of lessons will be provided to the student as soon as possible after the completion of either the planning or presentation of each lesson. The student will be permitted to offer any reaction or rebuttal to the evaluation which has been made in time for it to be of benefit to all concerned in the preparation of the next teaching lesson, or the full-responsibility demonstration, and
9. Each student will be permitted to offer any reaction or rebuttal to the final assessment made of his performance in the full-responsibility teaching demonstration before it is finally determined whether or not the student has adequately demonstrated teaching competence in terms of previously agreed upon standards.

Performance Standards for Competency Demonstration

Another aspect of the meaning of competency that had to be unraveled before progress could be made in the development of the assessment system was the matter of performance standards. This was a particularly troublesome concept for it was imbedded in both the nature of the competency to be demonstrated and the context in which it was to be demonstrated. For example, defining the objectives of instruction was a competency to be demonstrated, but there is nothing inherent in that competency descriptor that speaks to the quality expected (standard) in its performance. It also makes no reference to the context in which performance is to take place. This is equally troublesome since the performance standards for defining the objectives of instruction in the context of lesson teaching may be considerably different than in the context of short-term, full-responsibility teaching. Because of this interdependency of competency descriptor, the context in which a competency is to be demonstrated, and the performance standard set for its demonstration the task of becoming clear as to what the assessment system was to do and how it was to do it was more difficult than anticipated.

Another level of subtlety and complexity emerged in relation to performance standards as the assessment system developed. This was the distinction that had to be drawn between performance ratings and performance standards. As the system was planned initially it was anticipated that performance standards would apply to each competency that was being assessed. As the system evolved it was discovered that applying the concept of performance standards at that level of detail was simply not functional. Ratings of performance had to be applied at that level; i.e., at the level of each competency descriptor, but it turned out that performance standards seemed to apply best to performance within a particular demonstration context. Thus, as the system evolved during the first year of program operation, performance standards came to apply to performance patterns across competencies within particular demonstration contexts, rather than to individual competency demonstrations.

Such an arrangement in no way lessens the importance of individual competency assessments. These are still the central focus of the system, and their assessment provides the basis for much of the instruction that occurs in field settings. Their assessment also provides the basis for arriving at a judgment about performance standards, for these are defined in terms of individual competency assessments within a particular demonstration context. Seen in this way performance standards serve operationally as the exit requirements from a particular demonstration context, or entrance requirements to another. They also serve as the criterion measure for certification. The performance standards that were established during the first year of program operation for lesson teaching and short-term, full-responsibility teaching appear as attachment A.

The Approach Taken to Measurement

The approach taken to the measurement of individual teaching competencies was one of obtaining carefully delimited professional judgments, in the form of rating scale placements, as to the adequacy of a student's performance in a particular demonstration context. At least two separate professional judgments were obtained in relation to each competency demonstration, one from a student's college supervisor and one from

his school supervisor. An evaluative judgment was also obtained from a content specialist if a student requested it. The ratings were designed so as to accommodate the impact of acting differences on competency demonstration.

The rating scales assumed somewhat different forms for evaluating plans to evaluation presentations. In evaluating plans a three-point scale was used; in evaluating presentations, a five-point scale. This difference occurred because ratings were applied only to items in a plan that were unacceptable; i.e., if an item was unacceptable it had to be modified until it was. As a consequence the ratings applied to the plan were really only the upper three scale positions of the five-point scale applied to teaching presentations.

In addition to a judgment as to individual competencies, raters were asked to provide an overall judgment as to the quality of a plan or a presentation. These latter judgments were to be made only after all individual competencies had been assessed. To illustrate the nature of the overall rating system the forms used in the application of the scales to lesson teaching are appended as attachment B.

Two other features of the rating system are worthy of note, namely, the listing of items under each competency description that give focus to the description and the requirement that the behavioral or product indicators relied upon by a rater in responding to an item be recorded. In the evaluation of plans, each of the focusing items was rated (which turned out to be a poor practice). In the evaluation of presentations each focusing item was attended to, but only the competency was rated. This latter strategy proved to be reasonably functional, and provided the basis for an important revision of the forms for the coming year.

The decision to ask raters to record the indicators they relied upon in making a particular judgment grew out of the need to make the ratings as objective as possible. It was felt that the inclusion of the request to record indicator statements would help bring about objectivity in two ways. First, it would force raters to at least think about, and hopefully identify, the indicators used. Second, by compiling a listing of the indicators used by different raters in different settings, a guide on indicator use could be prepared that could be of value in training raters. Unfortunately, not enough care was directed to this aspect of the rating process during the first year of program operation, and a relatively spotted record of indicator use was the result. In the coming year the rating methodology has been changed in this regard, however, and indicator usage assumes a more prominent place within it.

Data Management and Utilization

Only the most rudimentary system was developed during the first year of program operation for the management and utilization of the data coming from the competency assessment system. A work-study student was responsible for the distribution of the various evaluation forms to students and college supervisors. Students were then responsible for getting the forms they received to their school supervisor, and for returning the completed school supervisor's forms to the college supervisor. A record of completed forms was maintained in each college supervisor's office for each of the students they were sponsoring, and a data summary sheet was prepared from each set of forms for use in a permanent file. The data recorded on these summary sheets were then put in computer storage for purposes of competency profile preparation and research. For illustrative purposes, the summary data form for lesson teaching appears as figure 1. The translation of this information to the computer permitted a series of methodological studies to be undertaken on the adequacy of the measures coming from the system (see the next section of the paper), and preliminary work was able to be done in the preparation of competency profiles. This task was sufficiently complex, however, that its completion was not possible during the first year of program operation.

Data management procedures that served online decision making in relation to lesson and short-term, full-responsibility teaching were left largely to the ingenuity of the college and school supervisors involved. Several guidelines to data use were provided, however. First, all conferencing with students about plans or performances in the classroom was to be data based, at least at point of departure. Second, a student had the right and obligation to quarrel with ratings or indicator statements whenever he or she felt they were inaccurate or unfair. Third, discussion of performance in the

[1] The assessment system were actually developed in support of the experimental program at OISE during its first year of operation. The first is the competency assessment system that is described in the present paper. The second is a program management system that is designed to systematically collect data on all aspects of program operation and make that data available at a time and in a form that facilitates program adaptation decisions. The program adaptation system is designed on the same principles as the competency assessment system (see pp. 9 to 14) and should therefore have the same degree of transportability.

RESEARCH SUMMARY: LESSON PREPARATION AND PRESENTATION

NOTE: The data summarized below are taken directly from LESSON PLAN and LESSON PRESENTATION evaluation forms. For some elements the data transfer is direct, that is, the element appearing on the present form has an exact counterpart on an evaluation form, and thus the rating that appears for the element here is taken directly from the evaluation form. In other cases, however, the data transfer requires manipulation, for example, further reduction or further summarization. Where this is the case, existing ratings are always used in the manipulations. NEW RATINGS ARE NEVER OBTAINED OR MADE FOR PURPOSES OF A RESEARCH SUMMARY.

Student's Name | Lesson Number | Person Who Provided The Rating

SELECTION A: LESSON PLAN

1. OBJECTIVES (average of the four ratings re objectives)
2. ADAPTING OBJECTIVES TO LEARNER CHARACTERISTICS (direct transfer)
3. SELECTING INSTRUCTIONAL MATERIALS AND PROCEDURES (average of four ratings re instructions)
4. EVALUATION (average of the three ratings re evaluation)
5. PLANNING NEXT STEPS (average of the two ratings re next steps)
6. THE FIT BETWEEN INSTRUCTIONAL LESSON AND CONTENT (average of the two ratings re fit)
7. SUMMARY EVALUATION OF THE LESSON PLAN (average of the ratings appearing above)

SELECTION B: LESSON PRESENTATION

1. OBJECTIVES (direct transfer)
2. ADAPTING THE LESSON TO LEARNER CHARACTERISTICS (direct transfer)
3. CLASSROOM MANAGEMENT (direct transfer)
4. INSTRUCTION (direct transfer)
5. EVALUATION (average of three summary ratings)
6. ACHIEVEMENT OF DESIRED LEARNING OUTCOMES (direct transfer)
7. PLANNING NEXT STEPS (direct transfer)
8. THE FIT BETWEEN INSTRUCTIONAL LESSON AND CONTENT (direct transfer)
9. SUMMARY EVALUATION OF THE LESSON PRESENTATION (average of the ratings appearing above)

Figure 1. Data summary form for lesson planning and preparation

classroom was to occur as close as possible in time to actual performance. Fourth, the college and school supervisor would confer as to the overall adequacy of the performance of a student before a student was permitted to advance to the next level of competence demonstration, even though the college supervisor was the person to make that judgment ultimately.

Within these broad guidelines all specific conditions of procedures and schedules were followed. No collection or synthesis of the data on the evaluation forms occurred before or after discussions with students, except for the summary that was prepared for permanent file and computer use. As the reader will see in the next section of the paper the failure to direct greater attention to the management and utilization of the data coming from the assessment of lesson and short-term teaching had its consequences.

Data management procedures were more exacting in relation to the student teaching (specialties) demonstration, even though much less work had been done developmentally on that demonstration context than the others. This undoubtedly reflected the fact that certification decisions were involved. The group, one that was developed to handle this level of decision required that a jury of college faculty and classroom teachers who had not served as the student's supervisors examine all data available on a student's performance and arrive at a decision concerning the match between performance and performance standards. The student's sponsor presents the data called for in the regard and provides the necessary context building that permits the jury to view the data in perspective. A segment of video tape showing the student teaching is included as a part of the student record.

This procedure was tested on [illegible] occasions during the first year of the program and found to be reasonably satisfactory. It needs to be pointed out, however, that how performance standards were not operating on these occasions and so our experience with these procedures did not represent a full test of them.

Data on System Operation

In order to determine how well the assessment system was working in the context of the experimental program, analyses were made of the various student applications of the system, the completeness of information coming from the system, the trustworthiness of that information, etc. Studies of this nature were carried out midway through the program and then again at its completion. The data reported in the paragraphs that follow are based on these analyses.

Of the [illegible] students that enrolled formally in the experimental program, all completed lesson teaching and [illegible] of them [illegible] and the performance standards set for that assessment [illegible] system contacts. Of the [illegible] students who completed [illegible] assessments, [illegible] completed [illegible] and [illegible] of them [illegible] and the performance standards set for that [illegible] contacts. Of [illegible] the data figures approximately [illegible] these [illegible] cases of the students who entered the program [illegible] [illegible] assessment experiences for one [illegible] from A. or [illegible] past faculty and [illegible] for one [illegible] student teaching. Five of the 56 students, however, dropped [illegible] [illegible] student teaching through the student teaching [illegible] [illegible] assessments, and three of these five [illegible] [illegible] component at the level of initial certification.

Failure to [illegible] to [illegible] data would suggest that the [illegible] system was working well. In order to [illegible] this is a fair judgment. A larger proportion of students [illegible] progress or even dropped from the experimental program [illegible] than is typically the case in the conventional [illegible] [illegible] of the program's college and school supervisors [illegible] also expressed that the instruments that they were able [illegible] to make of students' performance were [illegible] adequate and were satisfied that they had over time [illegible] [illegible] faculty also expressed that the data [illegible] [illegible] [illegible] to the college faculty facilitated instructional [illegible] [illegible] that accompanied the supervisory process.

Two [illegible] [illegible] sets of data sources from students in interacting [illegible] [illegible] with the system worked. The first is a set of data [illegible] that has to do with the appropriateness of performance [illegible] [illegible] and documentation by various evaluators. The [illegible] data are summarized in tables 1 and 2, and are [illegible] [illegible] on a number of counts. First, it is immediately [illegible] clear that the rating forms were not adequate in all [illegible] cases, and often many items applied [illegible] were not [illegible] or completely. Second, there are [illegible] [illegible] [illegible] comments, and more complaints that [illegible] [illegible] [illegible] forms. Third, the appropriateness of rating and documentation varied by class of users with the school [illegible] supervisors generally being the most conscientious [illegible] [illegible] filling out the forms completely. Finally, very few [illegible] [illegible] comments from the college faculty applied the [illegible] [illegible] to items in either ratings or documentation, and where they [illegible] [illegible] they were not very [illegible] [illegible] [illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]

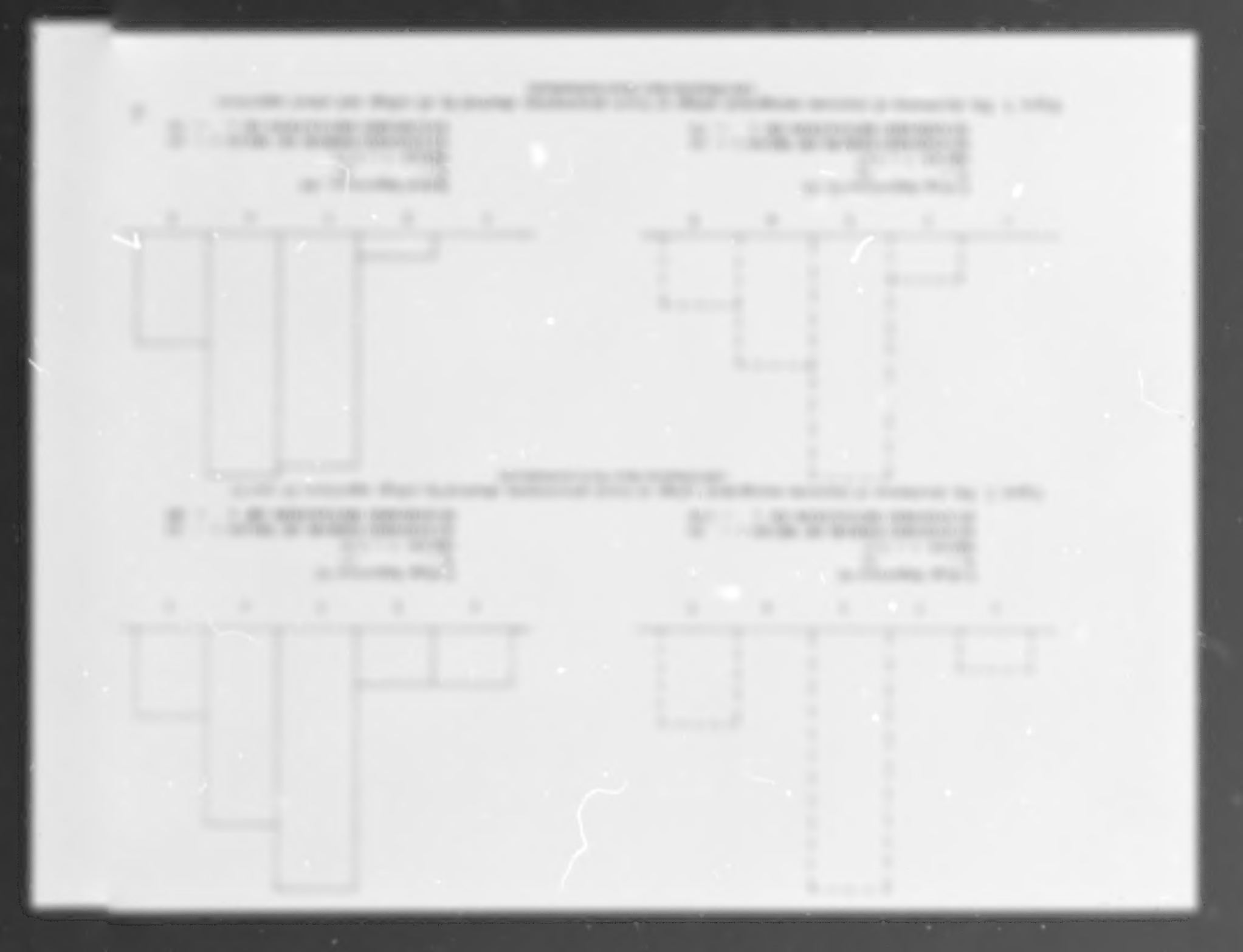
[illegible]

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

References

Schalock, H. D. and Garrison, J. H. The personalization of teacher education programs. In DeVault, et al. (ed.) "New Directions in Teacher Education: Problems and Prospects for the Decades Ahead." San Francisco: McCutchan Publishing Co., 1973. pp. 33-34.

Schalock, H. D. and Hale, J. R. "A Competency-Based, Field-Centered Systems Approach to Elementary Teacher Education. Volume I: Overview and Specifications. U.S. Office of Education, Bureau of Research, Project No. 89022. Portland, Oregon: The Northwest Regional Educational Research Laboratory, 1968.

Schalock, H. D., Kersh, B. Y., and Horyna, L. L. "A Plan for Managing the Development, Implementation, and Operation of a Model Elementary Teacher Education Program." Vols. I & II. Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1970.

Schalock, H. D., Kersh, B. Y., and Garrison, J. H. "Teacher Education in Transition: A History of CBTE at Oregon College of Education. Vol. I: The Experimental Program." In preparation.

Turner, R. L. "The Acquisition of Teaching Skills in Elementary School Settings." Bulletin of the School of Education, Indiana University, 1965.

Turner, R. L. Rationale for competency-based teacher education and certification: levels of criteria. In B. Rosner (ed.) "The Power of Competency-Based Teacher Education." Prepared by the Committee on National Program Priorities in Teacher Education. New York: Allyn and Bacon, 1973. pp. 3-8.

Attachment A. Performance Standards

PERFORMANCE STANDARDS FOR LESSON TEACHING

The first and simplest context within which teaching competency is to be demonstrated in the ETE program at OCE is lesson teaching. Competency at this level of teaching must be demonstrated before a student is free to engage in full-responsibility teaching. At least three lessons must be taught for purposes of formal evaluation, and standards must be met for both their preparation and their presentation.

Standards for Lesson Planning

In preparing for lesson teaching a reasonably detailed lesson plan must be prepared and reviewed by both college and school supervisors. The standards set for the preparation of plans are *item specific* standards, that is, both the college and school supervisor must indicate, independently, that every item to be attended to in the plan has been dealt with satisfactorily. This standard must be met before the lesson can be presented to children. If a plan does not meet this standard upon its initial review, it must be revised until it does.

Standards for Lesson Presentation

The standards set for performance in the presentation of lessons are *pattern* standards, that is, they apply to the pattern of performance demonstrated in the presentation of three or more lessons. Two standards are to be applied to the performance record of a student on the three or more lessons presented:

- —evidence of favorable performance on each of the teaching functions assessed in at least one of the three lessons presented;
- —evidence of favorable performance on the preponderance of teaching functions assessed in the three lessons presented. Preponderance is defined here to mean at least 75 percent of the functions assessed in the course of the three lessons presented will reflect evidence of favorable performance, and no more than 25 percent of the functions assessed will reflect evidence of unfavorable performance.

PERFORMANCE STANDARDS FOR SHORT-TERM (2-5 DAYS) FULL-RESPONSIBILITY TEACHING

When students in the ETE program meet performance requirements in lesson teaching they are free to enter the first full-responsibility teaching experience that is provided in the program. This is what is termed a short-term, full-responsibility teaching context. This experience requires a student in the program to assume full responsibility for planning and carrying out instruction in the schools for a minimum of 2 days and a maximum of 5.

Three kinds of standards are applied to a student's performance in short-term, full-responsibility teaching. Two of these correspond to the standards applied to lesson teaching. The third set of standards pertains to the utilization and management of affect.

Standards for Curriculum Planning

As in the case of individual lesson teaching, a curriculum plan for short-term, full-responsibility teaching must be approved before teaching can be undertaken. This requires that a reasonably detailed curriculum plan be prepared for a 2- to 5-day demonstration (as used here a curriculum plan consists of a number of individual lesson plans, and the relationships if any between them), the curriculum plan be reviewed by both college and school supervisors by lesson plan, and that both the college and school supervisors must indicate, independently, that each of the items to be attended to in the plan as a whole has been dealt with satisfactorily. In keeping with the generally more demanding requirements of the short-term, full-responsibility demonstration all lessons to be presented within the 2- to 5-day teaching period must meet acceptable standards before teaching can begin.

Standards for Curriculum Presentation

Standards for curriculum presentation in short-term, full-responsibility teaching are more demanding, and cover more aspects of teaching, than do the standards for individual lesson presentation. The first standard assumes the same form as one of the standards set for

the presentation of individual lessons. This is the pattern standard that requires evidence of favorable performance on the preponderance of teaching functions assessed in the sum of the lessons presented in the 2 to 5 days of full-responsibility teaching. The second standard for lesson presentation is also a pattern standard, and requires that over the course of the lessons presented in the period of full-responsibility teaching:

- —variety in learning activities will be provided;
- —variety in cognitive functions and levels in pupils will be exercised;
- —variety in affective expressions will be employed in teaching; and
- —positive feelings in pupils, such as excitement and interest, will be utilized in their learning.

Standards for Affect Management

The standards set for the management of affect in short-term, full-responsibility teaching take as their focus four dimensions of affective expression:

- —teacher responses to instances of pupil affect;
- —the management of pupil responses to instances of pupil affect;
- —the anticipation of pupil upsets and disruptions, and their redirection; and
- —the management of pupil upsets and disruptions when such occur.

Performance standards in relation to these dimensions of affective expression require that during the course of the full-responsibility teaching experience a student need only to perform effectively three of the four dimensions specified (any three will do), and that he or she needs to perform to this level on only one of the 2 or more days that he engages in full-time teaching. Such a standard reflects the view that the management of affect in a classroom is a complex matter, and that in an initial full-responsibility teaching situation, performance standards set for it should not be particularly demanding.

Attachment B

DEMONSTRATION CONTEXT: LESSON PREPARATION AND PRESENTATION

LESSON PLAN

(Attach to your Lesson Plan and Lesson Plan Evaluation Form)

Student's Name *Lesson Number*

RECORD OF NEGOTIATION

APPROVAL TO NEGOTIATE THE LESSON WITH A SCHOOL SUPERVISOR

College Supervisor *Date*

APPROVAL TO PREPARE A FORMAL PLAN FOR THE LESSON

School Supervisor *Date*

* * *

CONTEXT DESCRIPTION

STUDENTS TO BE TAUGHT

School

Grade

Number

Special Characteristics

CONTENT TO BE TAUGHT

Area

Expected Learning Outcomes

DATE(S) OF LESSON PRESENTATION

TIME(S) OF LESSON PRESENTATION

* * *

SUMMARY EVALUATION OF LESSON PLAN

(Obtain only after all elements of your plan have been evaluated)

I judge the plan as a whole to be of (circle one) →	ACCEPTABLE QUALITY	OUTSTANDING QUALITY
School Supervisor		
College Supervisor		
Content Specialist		

* * *

Field Test Format 2
Experimental Elementary Teacher Education Program
Oregon College of Education
December 1972

LESSON PLAN EVALUATION FORM

Student's Name _____ *Lesson Number* _____

Have the evaluators that check your plan initial each of the items listed that meets with their approval. If the treatment of an item is thought to be outstanding, have the evaluator draw a circle around his or her initials. Be sure to attach this sheet to your lesson plan.

ELEMENTS OF THE PLAN	*EVALUATORS OF THE PLAN*		
	School Supervisor	*College Supervisor*	*Content Specialist*
OBJECTIVES			
Are the learning outcomes expected from the lesson clearly stated?			
Are they appropriate and worthwhile outcomes, given the characteristics of the pupils to be taught?			
Are the indicators that are to be used as evidence of successful outcome achievement identified?			
Are the procedures to be used in obtaining evidence of outcome achievement identified?			
ADAPTING OBJECTIVES TO LEARNER CHARACTERISTICS			
Are there provisions for modifying the objectives of the lesson to meet individual pupil characteristics?			
SELECTING INSTRUCTIONAL MATERIALS AND PROCEDURES			
Are the instructional materials to be used in the lesson clearly identified?			
Are they appropriate to the learners to be taught and the learning outcomes to be achieved?			
Are the organizational and instructional procedures to be used in the lesson clearly identified?			
Are they appropriate to the learners to be taught and the learning outcomes to be achieved?			
EVALUATION			
Are there provisions for determining where pupils stand with respect to the desired learning outcomes of the lesson before it is presented?			
Are there provisions for feedback to pupils about their performance during the time the lesson is being presented?			
Are there provisions for determining where pupils stand with respect to the desired learning outcomes of the lesson after it has been presented?			
PLANNING NEXT STEPS			
Is there some indication in the plan of what would be done next with the pupils if the learning outcomes expected from the lesson *materialize*?			
Is there some indication in the plan of what would be done next with the pupils if the learning outcomes expected from the lesson *did not materialize*?			
MATCHING INSTRUCTOR, LESSON AND CONTEXT			
Does the lesson as planned appear to be feasible and appropriate to the school setting in which it is to be presented?			
Does the lesson as planned appear to be feasible and appropriate to the student who is to present it?			

DEMONSTRATION CONTEXT: LESSON PREPARATION AND PRESENTATION

SUMMARY EVALUATION OF LESSON PRESENTATION

Student's Name *Lesson Number* *Date*

We judge each of the lesson elements that follow to be of (circle one) →	*UNACCEPTABLE QUALITY*	*ACCEPTABLE QUALITY*	*OUTSTANDING QUALITY*
OBJECTIVES			
ADAPTING THE LESSON TO LEARNER CHARACTERISTICS			
QUALITY OF PRESENTATION			
Classroom Management (Transitions, Termination, and Student Attention)			
Instruction (Materials, Procedures, and Organization)			
EVALUATION			
Pre-assessment			
Feedback during the lesson			
Feedback after the lesson			
Achievement of desired learning outcomes			
PLANNING NEXT STEPS			
THE FIT BETWEEN INSTRUCTOR, LESSON, AND CONTEXT			

School Supervisor *College Supervisor*

Student Presenter

Content Specialist *Student Recorder*

* * *

Field Test Format 2
Experimental Elementary Teacher Education Program
Oregon College of Education
December 1972

DEMONSTRATION CONTEXT: LESSON PREPARATION AND PRESENTATION

LESSON PRESENTATION EVALUATION FORM

Student's Name: _____ Content Area: _____ Date: _____ School or College Supervisor: _____

Lesson Elements	Evaluator's Judgment (check one)			If a Behavior Of The Student and or Pupils That Led You To Make The Judgment You Have Made	Suggestions For Improvement
	Yes	No	Insufficient Evidence To Judge		
OBJECTIVES					
Were the learning outcomes expected from the lesson clearly explained to the pupils?					
Were the learning outcomes expected from the lesson understood by the pupils?					
Did the pupils view the expected learning outcomes as appropriate and worthwhile?					
ADAPTING THE LESSON TO LEARNER CHARACTERISTICS					
Was the lesson carried out in a manner that reflected adaptations					
Was the lesson carried out in a manner that reflected adaptations to pupil characteristics?					
QUALITY OF PRESENTATION					
Was the transition between the previous learning activity and the present lesson effective?					
Was the pace of the pupils maintained throughout the lesson?					
Were instructional materials, resources, and organizational strategies used appropriately?					
Was the conclusion (set "wraps up" of the lesson effectively handled?					

COMPETENCY DEMONSTRATION CONTEXT: LESSON TEACHING

COMPETENCY CLUSTER IV, DISPLAYING PUPIL ACHIEVEMENT
COMPETENCY INDICATORS: PRODUCTS OF A STUDENT'S BEHAVIOR

Student's Name | *Lesson Number* | *Date Presented* | *Time Presented* | *Supervisor of Content Specialist's Name*

DIRECTIONS TO EVALUATORS. When evaluating a lesson presentation rate each of the achievement display functions (bold face items) that appear on this form. All ratings are to be entered in the boxes provided. RATINGS FOR ALL ITEMS THAT APPEAR ON THIS FORM ARE TO BE BASED UPON PRODUCTS OF A STUDENT'S BEHAVIOR, RATHER THAN UPON BEHAVIOR PER SE. These ratings are to reflect one of five possible judgments:

(1) **UNACCEPTABLE**

The products of the student's effort to display learning outcomes related to the lesson reflect an *insufficient* level of thoroughness, inventiveness, and care in relation to this teaching function.

(2)

(3) **ACCEPTABLE**

The products of the student's effort to display learning outcomes related to the lesson reflect an *adequate* level of thoroughness, inventiveness, and care in relation to this teaching function.

(4)

(5) **SUPERIOR**

The products of the student's effort to display learning outcomes related to the lesson reflect an *exceptional* level of thoroughness, inventiveness, and care in relation to this teaching function.

ACHIEVEMENT DISPLAY FUNCTION	*SAMPLE INDICATORS*	*INDICATORS THAT SERVED AS A BASIS FOR JUDGMENT*
DISPLAYING PRELESSON AND POSTLESSON ACHIEVEMENT ☐ Were the results of the assessment before the lesson recorded and displayed in such a way that they were easily interpreted? (Please attach the display if one has been prepared; sample display forms are found in the "OCE Guide to Competency Assessment in Lesson Teaching") Were the results of the assessment after the lesson recorded and displayed in such a way that they were easily interpreted? (Please attach the display if one has been prepared; sample display forms are found in the "OCE Guide to Competency Assessment in Lesson Teaching")	Positive to Negative Examples • Results of at least the postlesson assessments of learning outcomes are recorded and displayed so that anyone looking at the information can easily determine where each pupil stands with respect to the learning outcomes expected from the lesson • Postlesson learning scores are summarized to show the proportion of pupils who achieved the learning outcomes expected from the lesson, and the proportion that did not • Prelesson and postlesson assessment information is not recorded or displayed • Postlesson achievement data are displayed but are not summarized to reflect the proportion of pupils who achieve the learning outcomes expected from the lesson, or the proportion of students who fail to do so	COMMENT

Estimating Costs of a Competency-Orientation

By

BRUCE R. JOYCE

How do we build a competency-oriented certification and training system? Personally, I could never go through all the necessary complicated steps, although there are people here who could. I'm much more interested in trying to find out what teaching skills will pay off, whether they have to be idiosyncratic, whether anybody can even train us all in the same ones. It's conceivable that we all may not be capable of the same ones; maybe in the long run we will have to do our own thing as teacher competence, in the end, may turn out to be personal—or it may be external and trainable. Also, teaching is elusive. You can measure 2,000 reasonable-sounding competencies and not have caught the essence of teaching. So far the research hasn't validated many competencies, but I'm not prepared to give it up.

To determine competence, we also have to face some rough questions. For example, we're not even sure that as teachers we should be playing the roles we do. I keep critizing the role of the self-contained classroom teacher mostly because it nearly killed me. I had to go into college teaching to keep alive, but I simply couldn't teach six things to 36 kids who were that different from each other. I didn't have the capacity for the job, and I had to become a professor in order to survive. I'm not sure that I want to spend any time finding competencies for the present roles teachers play. We might do much better to design better schools and find out what competence we need to make them work.

I know nothing about costs, OK? I couldn't cost-account a major program, but last summer, some of the New York State Education Department people asked me to help them estimate how much a competency program would cost to develop. I agreed to do it, and I wrote them the following paper.

Largely because of the massive research and development effort entailed if competency-based training and certification are to be achieved, it has been suggested that development and implementation be centralized. An alternative to accommodate local needs and bring about a broad base of participation in development is that the competency-orientation should be required and relatively small local consortia (probably made up in each case of a few school districts and colleges) should determine competency criteria, develop programs, and implement the programs and the certification prerequisites. The latter option is attractive because it promises to involve so many people and to insure relevance to local school needs.

The centrally mandated system is unacceptable on both political and substantive grounds, politically because it would give considerable power to a few persons (competency-oriented certification and training is much more likely to affect who will actually be permitted to teach than course-based certification and training, however centralized), and substantively because not enough is yet known about the identification of teaching competencies to permit anyone to develop a widely mandated set of competencies for any particular category of teachers with any great confidence that those competencies will stand up.

Centering development around many local centers is acceptable as a process, but is not without serious problems. The more local centers there are, the more difficult it will be to organize teams of sufficient expertise to develop really strong training and certification systems. Major universities, working with strong school districts, have had trouble doing this. Yet, strong development teams are necessary. High quality in a competency-orientation is essential because both certification and training procedures will be more powerful and mistakes (such as emphasizing trivial competencies) will be magnified.

Many districts and smaller colleges are unaccustomed to freeing personnel for development, but a sustained effort by a large team is essential if the task of identifying competencies, organizing training and assessment, and implementing certification procedures is to be done effectively.

Without coordination and quality control, broad participation in the program development could be an expensive and ineffective process.

Purpose

In this paper an attempt is made to identify the tasks of creating a competency-based teacher training and certification system and to estimate the costs of completing the tasks adequately. In addition, a basis is provided for estimating the effect on costs of several possible options for organizing the process of development and implementation. This can be used as a basis for estimating cost when several possible options for organizing the process include broad local input into the system. To estimate basic technical costs, an estimate is made of the cost of development by a compact, independent team of experts who would simply accomplish the technical tasks of producing a prototype training-certification system. This option (referred to as process option A1) is not a realistic choice, but yields a fairly reliable cost estimate from the technical point of view, which permits an estimate of the cost if several possible organizational processes were employed (e.g., strictly local development (process B1), local development with central technical assistance and development services (process C), and a central state system with provision for local options (process D)).

Sources of Cost Estimates

How do we arrive at the costs of the tasks and subsequently, the costs of the process options? There are three sources of our estimates. One is the experience of the Bureau of Research Teacher Education Project, especially the costing procedures included in their feasibility studies. The second is the cost of developed performance-based materials such as those in "Materials for Modules" (appendix B). These two sources can help us determine the base costs represented in process option A (the least desirable process option in terms of creating a statewide training-certification system). The third source is the least reliable and represents estimates of the probable increase in the base cost due to the greater complexity and duplication of effort of the other process options.

Technical Tasks and Process Options

Certification and training are interrelated parts of the same system which serves as the basis for preservice teacher education, inservice or post education, and the certifying and diagnostic procedures related to both of these.

The purpose of the development of competency-oriented systems is to unite preservice education, initial certification, inservice education, and continuing or extended certification around a system of teaching competencies which form the goals of preservice education, the standards for initial and extended licenses, and the basis for diagnosis of performance of inservice teachers and prescription of inservice training. It is probably not possible or desirable to separate licensing or training at the preservice or inservice levels under the competency orientation.

The entire system for competency-oriented certification and training depends on the creation of base interrelated storage and retrieval systems plus the organizational and communications networks necessary to create and implement them. The creation of these systems and networks constitutes the technical side of the development process. The creation and management of the organization necessary to create the systems constitutes the process side. Both technical and process sides present options which affect costs greatly.

The Technical Side

The technical side of competency-based certification and training involves the creation and validation of base systems. These are as follows:

A storage and retrieval system of the specifications of teacher competency.

Since teaching is complex the number of competencies which are likely to be specified is large indeed. The Bureau of Research Models averaged between 2,000 and 3,000 competency specifications and these represented efforts by single unified institutions or small consortia rather than statewide consortia of diverse institutions and representation. It seems reasonable to suppose that the number of competencies will increase as the political base for establishing them is broadened. Hence, when teacher associations, representatives of school administrations, the public and students all contribute, as well as expert teams from universities and state departments of education, the number of potential competencies of a teacher will be large and the process of identifying the most important competencies may be complex.

A storage and retrieval system of mediated instructional systems and unmediated components designed to produce the competencies.

This storage and retrieval system represents the means of teacher education. The number of instructional systems will be equal to or somewhat greater than the number of competencies which are specified. Also, the convenience of each instructional system is much greater than the specification of any competency. Thus, the production of software to fill this system will be an extensive task. For many people, it will be difficult to specify critical competencies, but developing the program elements to achieve them will be even more complex.

A system of assessment devices designed to determine the effects of the instructional systems, and competency-mediated components and to measure the competencies specified in system A.

Without an assessment system competency-based certification would be impossible as would be the assessment of the effectiveness of program elements. Properly organized, the assessment system provides a diagnostic profile of the teacher candidate and provides also the means for tracking his progress to determine when certification should be granted. In the case of in-service teachers, it provides the means for diagnosing the state of their competency and relating them to the instructional systems to be used in inservice education.

A management system for coordinating subsystems 1, 2, 3, and 4 above.

The magnitude of the three other systems requires the use of a contemporary management system for diagnosis, prescription, tracking and progress, and providing feedback to teacher candidates, teachers, and program and licensing managers. Without a complex automated management system individualization or personalization would be impossible in a program and it would be implementation, for the complexity of specifications and training devices is so great as to defeat any presently available option for program control. Without a modern information system there would surely ensue at the point of implementation. It is possible to imagine a statewide management system which individual teacher training institutes would use to participate in the identification of competencies, and they might use a central system to withdraw instructional systems and assessment devices for their particular training programs.

The research and development effort needed to produce the four instructional systems is enormous. They represent the core of the substance of competency-based teacher education and licensing.

The Process Side

If it is desired to have participation by all relevant groups in the creation of the competency-based teacher education and certification system then suitable communications networks and organizations have to be set up to permit teachers, students, teacher organizations, representatives of the public, subject matter experts, experts on teaching, and representatives of state departments and colleges to participate in competency identification, the selection of instructional systems, and the creation of the assessment systems. The process of creation can reasonably be divided into three levels: specification of the system, development, and implementation. It is possible to imagine a state or statewide network providing participation by all relevant groups in the specification level of the creation of the basic system. Development will probably have to be organized at a few major centers, funded to bring together talent to create the operating systems. It is possible to imagine a network stretching across the state so that many local units might contribute something to the development. However, development is an exceedingly complex task, especially if the products are to be tested as they are developed. Thus, a few major centers probably will do most of the actual production of the instructional systems and testing devices. However, broad-based groups can evaluate the products, and a representative group can monitor the entire operation.

Minimum, Lateral, Statewide Competencies and Extended Local Competencies

It is possible to use some of the critical competency specifications for minimum statewide certification—they could represent minimum standards for preservice and inservice training. A much larger number of specifications might be prepared representing local needs. These would be used by individual teacher training centers for training purposes and for determining local certification in addition to state certification. Imagine, for example, a network of local centers representing the local school districts, higher education institutions, community representatives, teacher associations, etc. Suppose that there are about 20 of these in various parts of the state. They produce competency specifications and agree on those which will be used for minimum specifications for any given type of teacher (elementary, secondary, special education, etc.). Each of the local centers, however, would produce specifications which were deemed important by

the people in that area that unless more and enhanced life experiences opportunities are created for individuals participation. These additional specifications were likely to center on the specific given energy circle and instructional systems would be designed to address these and associated concerns as issues arise. When a complete statewide system of K-12 credits costs of local needs would have to be much larger than the system which would produce only some improvements. Clearly the greater the participation of interested parties and the more numerous the local options, the more costly the creation of the system. From the point of view of organization and communication and also from the point of view of the use of learning across participation and with provision for local options costs will be greater.

The cost would be nothing for the cost of giving each local unit or center including several local units to create their own competency-based teacher education and certification system. The cost of helping and training competency-based programs with local options will be high as will the cost that is presently the minimum which would be unacceptable. It would cost $1 or $2 million for each local unit to create its own system. What if there were 10 systems in the state could developed by local units. The cost would be in the neighborhood of $10 to $15 million. It would seem to be more reasonable to use in order to create a statewide system that had provisions for special local needs and interests.

Incremental Costs and Process Options

The four costs involved in the creation of a competency-based program, plus the development of a certification system, represented the five costs required to activate a statewide or comprehensive training and certification system. Nearly one of each of these five costs, considerably understanding on the system options which are selected. As table 1.1 lists process options and the general cost for the five systems.

Process A (acknowledging the competency principle) is the least costly option in order to achieve the low the disadvantages. Due to statewide acceptance would not be high as all the methods.

Process B (acknowledgement of competency system for local and local systems) would be very costly. The cost for each complete system would be higher than that of a system developed by an "expert" could be estimated that the cost would be $4-5 million dollars and this would be multiplied by the number of local systems.

Process C (local or development supported by a few development centers) would be considerably less costly than B because the development systems would reduce duplication of effort.

Process D (acknowledgement of a statewide system with local options) would be much higher than B but much lower than E or F. As with former could be an acceptable cost-wide process option plan. It appears to be the best cost option.

Process E is low in costs because there continues to be partial duplication but no cost estimates because it has a low cost which can be duplicated by future implementing the assessment of competency and duplication of all four suggested in the set other options.

Process F results in no system comparable systems. It gives local needs the no further play, but the use of system costs escalates as the cost of each system or that each local system economy by cost expenses and the cost

Table 1 Cost Factors of Incidental Costs by Process Options

Process	A. Base Cost for Higher Costs	B. Local Consortia or Areas	C. Each of 5 systems (Plus Assessment Center & Consortia)	D. Base with Service Work of Local Systems
1. Competency System	1	1.5 one set of consortia	1.25 one set of consortia	1
2. Instructional System	1	1.5	1.25	1
3. Assessment System	1	1.5	1.25	1
4. Management System	1	1.5	1.25	1
5. Certification System	1	1.5	1.25	1

Bibliography

[1] Abadi, A. et al. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 308–318. https://doi.org/10.1145/2976749.2978318

[2] Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 308–318. https://doi.org/10.1145/2976749.2978318

[3] Abdalla, M., Bellare, M., & Rogaway, P. (2001). The oracle Diffie-Hellman assumptions and an analysis of DHIES. *Topics in Cryptology — CT-RSA 2001*, 143–158. https://doi.org/10.1007/3-540-45353-9_12

[4] Abdalla, M., Bellare, M., & Rogaway, P. (2001). The oracle Diffie-Hellman assumptions and an analysis of DHIES. *Topics in Cryptology — CT-RSA 2001*, 143–158. https://doi.org/10.1007/3-540-45353-9_12

[5] Abdalla, M., Bellare, M., & Rogaway, P. (2001). The oracle Diffie-Hellman assumptions and an analysis of DHIES. *Topics in Cryptology — CT-RSA 2001*, 143–158. https://doi.org/10.1007/3-540-45353-9_12

[6] Abdalla, M., Bellare, M., & Rogaway, P. (2001). The oracle Diffie-Hellman assumptions and an analysis of DHIES. *Topics in Cryptology — CT-RSA 2001*, 143–158. https://doi.org/10.1007/3-540-45353-9_12

[7] Abdalla, M., Bellare, M., & Rogaway, P. (2001). The oracle Diffie-Hellman assumptions and an analysis of DHIES. *Topics in Cryptology — CT-RSA 2001*, 143–158. https://doi.org/10.1007/3-540-45353-9_12

[8] Abdalla, M., Bellare, M., & Rogaway, P. (2001). The oracle Diffie-Hellman assumptions and an analysis of DHIES. *Topics in Cryptology — CT-RSA 2001*, 143–158. https://doi.org/10.1007/3-540-45353-9_12

[9] Abdalla, M., Bellare, M., & Rogaway, P. (2001). The oracle Diffie-Hellman assumptions and an analysis of DHIES. *Topics in Cryptology — CT-RSA 2001*, 143–158. https://doi.org/10.1007/3-540-45353-9_12

[10] Abdalla, M., Bellare, M., & Rogaway, P. (2001). The oracle Diffie-Hellman assumptions and an analysis of DHIES. *Topics in Cryptology — CT-RSA 2001*, 143–158. https://doi.org/10.1007/3-540-45353-9_12

University of Wisconsin, Elementary Teacher Education Project (OE-58025, four volumes), Washington, D.C.: U.S. Office of Education Bureau of Research, 1969.

II. Feasibility Studies

The Florida State University, A Feasibility Study of the Florida State University Model for the Preparation of Elementary School Teachers, Washington, D.C.: U.S. Office of Education Bureau of Research, 1969.

Michigan State University, Feasibility Study: Behavioral Science Teacher Education Program, Washington, D.C.: U.S. Office of Education Bureau of Research, 1970.

Ohio Consortium, The Feasibility of Educational Specifications for the Ohio Comprehensive Elementary Teacher Education Program (Phase II), Washington, D.C.: U.S. Office of Education Bureau of Research, 1970.

Oregon College of Education, A Plan for Managing the Development, Implementation and Operation of a Model Elementary Teacher Education Program, (three volumes); Washington, D.C.: U.S. Office of Education Bureau of Research, 1970.

Syracuse University Protocooperative, A Study of the Feasibility of the Refined Syracuse University Specifications for a Comprehensive Undergraduate and Inservice Teacher Education Program for Elementary Teachers, Washington, D.C.: U.S. Office of Education Bureau of Research, 1969.

University of Georgia, The Feasibility of the Georgia Educational Model for Teacher Preparation—Elementary, (three volumes), Washington, D.C.: U.S. Office of Education Bureau of Research, 1970.

University of Massachusetts, A Feasibility Study on the Model Elementary Teacher Education Program (Phase II), Washington, D.C.: U.S. Office of Education Bureau of Research, 1970.

The University of Toledo, The Feasibility of Educational Specifications for the Ohio Comprehensive Elementary Teacher Education Program (Phase II), Washington, D.C.: U.S. Office of Education Bureau of Research, 1969.

University of Wisconsin, Wisconsin Elementary Teacher Education Project Feasibility Study, (two volumes), Washington, D.C.: U.S. Office of Education Bureau of Research, 1970.

III. Other Sources

Cruikshank, Donald R., et al., The Ohio State University Analysis of the Nine Comprehensive Elementary Teacher Education Models (CETEM), Washington, D.C.: U.S. Office of Education, National Center for Educational Research and Development, 1970.

Engbretson, William E., Analysis and Evaluation of Plans for Comprehensive Teacher Education Models, Washington, D.C.: U.S. Office of Education, 1968.

Joyce, Bruce R., "Variations on a Systems Theme: Comprehensive Reform in Teacher Education," Interchange, Vol. 1, No. 3 (1970), pp. 83-95.

Klatt, Judith and Walt Le Baron, System Development Corporation, A Short Summary of Ten Model Teacher Education Programs, Washington, D.C.: U.S. Office of Education, National Center for Educational Research and Development, 1969.

Le Baron, Walt, System Development Corporation, Systems Analysis and Learning Systems in the Development of Elementary Teacher Education Models, Washington, D.C.: U.S. Office of Education, National Center for Educational Research and Development, 1969.

Proceedings of 22d Annual Georgia Teacher Education Conference: TEACHER EDUCATION MODEL, Journal of Research and Development in Education, Vol. 2, No. 3 (Spring 1969).

A Reader's Guide to the Comprehensive Models for Preparing Elementary Teachers, Washington, D.C.: ERIC Clearinghouse on Teacher Education and the American Association of Colleges for Teacher Education, 1969.

System Development Corporation, Analytic Summaries of Specifications for Model Teacher Education Programs, Washington, D.C.: U.S. Office of Education, National Center for Educational Research and Development, 1969.

Funding Competency-Based Teacher Education

By

HERBERT HITE

Competency-based teacher education costs differ from the costs of standard teacher education because the two programs differ. There are differences in the required performances of students and also in the roles of faculty. Program materials are different and so are assessment procedures. The sources of funds for CBTE, however, are not different from the sources for traditional, or standard, programs. The same sources of revenue for higher education which are always in short supply must implement the CBTE programs. Therefore, the analysis of costs for CBTE is most meaningful when these costs are expressed in terms of funding formulas for other higher education programs.

Among the program variables of CBTE which are reflected in differing costs from standard programs are the following: [1]

1. Instructional materials, units, and techniques are individualized. If the criterion of CBTE is the success of the individual student of teaching, then CBTE materials must allow for different rates and different options by these individual learners. These materials constitute a higher cost factor than standard program materials.
2. For the most part, the CBTE faculty work in a one-to-one relationship with students—as counselors, evaluators, explainers, analysts. The changed role of faculty has a higher cost than faculty in the traditional lecturer role.
3. The process of admission to CBTE appears to be a more complicated and hence a more costly process than admission to standard programs. Prerequisites for success, which should constitute admission criteria, seem to be more difficult to assess than the usual examination of students' grade-point averages.
4. Complex management systems may be a significant cost item.
5. Students in the CBTE programs are usually full-time interns or observers or work over varying periods of time with individualized study materials. Full-time students rather than number of classes are the bases for measuring costs. This is not necessarily a higher cost factor, just different.
6. In CBTE, school personnel have a significantly larger role in the instructional process and the cost of their participation is relatively higher.
7. More instructional time is involved in assessment of student progress, and this process is usually more complex than grading in courses. Evaluation costs are usually higher.
8. Decision making regarding all phases of the CBTE program involves not only college personnel but also school and community members. The large amounts of these people's time which is needed are a significant cost factor, and no funds have been available for this function in the past.
9. Specialist or consultant services to assist state agencies in program approval may result in additional costs for CBTE.
10. Competency-based programs leading to advanced certification will be centered in school districts, and the development and administration of these new programs will be an additional cost factor eventually.

In a study of costs for the new CBTE certification standards in the state of Washington, which was requested by the state's legislature, two CBTE programs were analyzed.[2] The study suggests that for the time being, the following factors are additional costs of CBTE relative to standard teacher education programs in that state:

1. Add 50 percent to the costs of the standard teacher education to account for the variable of individualized mode of instruction.
2. An additional 50 percent will be required for developing each new CBTE program.
3. An additional 50 percent will be needed to pay for released time of school personnel involved in consortium arrangements—policy making, program planning, securing program evaluation, etc.

[1] These 10 factors are described in a paper given for the Teacher Corps in June 1972. "The Economics of Competency-Based Teacher Education," Herbert Hite, Research Foundation of the State of New York, Albany, N. Y., 1972.

[2] One of the programs analyzed in the cost study was described in the paper previously cited.

The recommendations to the Washington State Legislature are for the colleges and universities to receive double funding for new CBTE programs and the school districts to receive funds for consortium activities.

Some of the costs for CBTE are the same if the program includes one student or 1,000 students. These costs include the development of policy councils, the certification process, the development of basic competencies and assessment processes, and a basic set of individualized materials. Some costs vary according to the number of students: the number of faculty per student, the number of laboratories or public school classrooms, the number of sets of materials. Some costs represent quality judgments, such as the number of options to be made available to students, the ratio of faculty to students, the degree of refinement of structional modules and assessment procedures.

Obviously, a good deal of the difference in costs of CBTE and standard programs depends upon decisions as to the nature of specific CBTE programs. For example, the definition of competency can make a difference in program costs. If competency is defined as a set of teaching performances, then the focus will be upon highly refined and tested instructional devices to elicit teaching behaviors. If the emphasis is upon products of teaching; i.e., pupil achievements, then the emphasis will be upon the field setting for working with learners and evaluating the outcomes of this activity.

The CBTE model at Western Washington State College is based upon product criteria. The candidate is required to demonstrate competency by bringing about "appropriate" changes in the behavior of elementary or secondary pupils. The major components of the model are an entry phase, in which the candidate meets admission criteria, a knowledge phase in which he acquires a repertoire of information about the substance of what he will teach, a laboratory in which he acquires and demonstrates knowledge and skill about teaching and learning principles and applies these skills, and a practicum in which he demonstrates his ability to apply principles of teaching and learning in a variety of situations. The laboratory and practicum constitute two quarters of full-time study and practice in a teaching center.

The program depends upon close working relationships in the laboratory and practicum among the students, the professors, and cooperating teachers who go through a special training program to prepare as field supervisors. The basic essentials of the program are evaluations of entry into the program, diagnosis of knowledge necessary to succeed in the laboratory, a test of instructional competency as an entry into the practicum and modified T.I.C.'s as evidence of mastery of the basic model and hence certification. Study materials in the laboratory consist of 52 modules and assorted readings, films, filmstrips, tapes, observation assignments, and mini-teaching arrangements. The students design their own study programs leading to the T.I.C., which is the basic measure of competency, by negotiating learning contracts with the clinical professors. Study contracts vary.

After 4 years of partial success and occasional failures, the clinical faculty has arrived at a working model. The costs of this working model are heavier than the costs for standard teacher education at Western. The cost studies which have been completed suggest that even after initial years of program development have been completed, there are these factors to add to the costs of the standard programs at Western:

1. About 15 percent additional is needed to fund extra support costs, such as visits by campus-based faculty to the teaching centers, and special individualized learning modules.
2. About 25 percent additional is needed for the continuous development of the program components—evaluation instruments, adaptations of modules, processes for orienting experienced teachers to the model, etc.
3. About 25 percent is needed for the administration of each teaching center, including negotiations with school personnel concerning consortium arrangements.

This is a modest proposal for extra funding compared to other suggestions for funding CBTE. The Western model is still developing, but it is based upon modest resources. Compared to the Elementary Teacher Education models, funded in 1968 by USOE, this is a Model T compared to a brand new Lincoln Continental. The present program does produce the required demonstrations of competency, however, and there are considerable affective gains over the standard program in the views of faculty and students.

Still, even a modest program needs somewhere between a 50 and a 100 percent increase in funding for teacher education. Where does the funding come from? In the real world, deans of education do not confront the legislature or the boards of regents with demands for so many extra millions to fund a new program. Colleges and universities have only a few alternatives.

These are the possible alternatives a college has for finding the additional resources for CBTE:

BEST COPY AVAILABLE

1. The college might maintain the existing level of faculty and instructional support funding but limit the enrollment of students in the program.
2. The college might require more time or credit hours by students who were candidates for a certificate under the CBTE plan, thus increasing either tuitions or student credit hours and consequent funding.
3. The faculty might seek an adjustment in the formula for support of programs. Field-based or laboratory programs might be funded at twice the level of campus-based programs.
4. The college faculty might seek outside funding —grants or fellowships.
5. Schools might contribute by sharing the cost through support of a complementary inservice CBTE program.
6. Colleges might develop graduate programs, with their higher rates of funding, as adjuncts to the competency-based undergraduate program.

In Washington, master's level instruction in state institutions is funded at 2.3 times the level of upper-division student instruction. Clearly this level of funding would provide what the faculty at Western have defined as their need for CBTE programs. This adjustment would require action by the state legislature or a major and radical change in programing teacher education within the college. Neither possibility seems likely.

Outside funding is possible, but not very dependable as a base for a continuing extra program cost. The college could reduce enrollment and maintain the existing level of faculty and support programs, but this is politically impossible at Western, although this may be possible at other institutions. A few school districts will share costs. In fact, one school district reportedly has a budget of $150,000 for advanced certificate programs under the CBTE model. These school funds, however, are usually restricted to payments to school personnel for released time. This leaves the college administrator where he always is—with the same old system to try to manipulate. Basically the game is to generate sufficient credit hours to secure the necessary entitlement for faculty and instructional support.

The CBTE program at Western operates in four teaching centers. Each center consists of a team of clinical professors, students, and cooperating teachers. In a center with four clinical professors, here is a strategy for generating the necessary student credit hours and getting the necessary work accomplished.

Two clinical professors will have eight to 10 new students each quarter. As the students remain for two quarters, this means that the instructional load for each of the two professors will be 32 full-time students, at the undergraduate level. These two professors between them generate three and one-third full-time faculty equivalent entitlements.

Herbert Hite

One clinical professor takes on half as many students, or a total of nine, which is about two-thirds of a full-time equivalent faculty load. He is also responsible for the administrative matters of the center, including negotiations with school personnel.

One clinical professor will direct 25 part-time graduate students in a four-credit practicum. The students are teachers and volunteer to act as cooperating adjunct faculty in the CBTE program. The credits they earn may apply to an advanced certificate. Their practicum is concerned with development of the processes for working with students in the program, including evaluation criteria and techniques. The major responsibility of this clinical professor is program development. The 25 students in the four-credit graduate practicum generate one full-time equivalent faculty load.

The four clinical professors, then, generate five CBTE units, which entitles the center to the needed instructional materials and services—which are additional to those required in campus-based programs. Also, two of the professors earn student credit hours in the process of administering and developing the CBTE program.

The four clinical professors may modify their own assignments. They are jointly responsible for 45 full-

time undergraduates and 20 part-time graduates, however they divide the responsibilities of the teaching center. These figures are only illustrative.

A variation on this model is a center in which the undergraduate intern, in his final quarter of demonstrating competency, relieves an experienced teacher, who becomes a full-time master's candidate. The intern is backed up by a first-quarter laboratory student who acts as a teaching assistant part time. Both are closely supervised by the clinical professor. In this model, instead of supervising 10 full-time undergraduates, the clinical professor is responsible for four full-time graduate students, four interns, and four laboratory students, in four elementary [illegible].

These figures and load descriptions represent a few of the ways in which the CBTE faculty may generate the amount of credit hours needed to support the costs of CBTE which are above the costs of standard programs. The experience at Western indicates that these kinds of loads are feasible, given the nature of the CBTE program which has been developed at that institution.

Summary

Competency-based teacher education is more expensive than the standard programs. The variables within CBTE programs which affect costs are many, and, additionally, there are many kinds of value judgments about programs which will affect the level of funding needed for a given CBTE program. Funding for a models approach to CBTE, which is largely field-based, is possible within the usual formulas for generating support for higher education. The experience of clinical professors indicates that the credit hours generated by full-time undergraduate students and part-time graduate students (cooperating teachers) are sufficient to fund CBTE. State officials could simplify the funding problem, however, by supporting CBTE as a "high-cost" college program, equivalent to graduate study or to study in the other professional fields such as health sciences or law.

It seems clear from initial experiences with CBTE that costs are greater than with regular programs. How much greater depends upon a number of judgments which are made by the CBTE program designers.

The high cost factor which is common to all CBTE programs is the individualized mode of instruction, which is necessary if teacher candidates are to demonstrate competency criteria. Other expenditures, particularly expenditures for program development and for management, can vary wildly from CBTE program to CBTE program.

The cost decision may hinge upon the way the faculty define competency. If competency is defined in terms of specific teaching behaviors which are assumed to be related to changes in pupils, then the modules and other training materials will need to be refined to the point that it is possible to predict a given level of performance on the part of the consumer of those modules. It follows from that decision that teaching modules and strategies must be highly refined. Costs can be very high to achieve this end.

Competency may be defined, however, as changes in the performance of the teaching candidate's clients. In that case the specific modules or strategies will be effective inasmuch as they assist an individual candidate to implement growth on the part of the candidate's pupils. The emphasis in program development will then be upon the field setting, the education of cooperating school personnel, the evaluation of the products of teaching. The costs for these components of CBTE are more likely to be absorbed within the ongoing school program for pupils and the inservice education for teachers.

If the CBTE faculty choose to design all their own learning modules, the costs will be relatively high—they could be astronomical. The cost factor seems comparable to the costs of developing programs of instruction or instructional systems. There are choices to be made concerning the quality levels of the CBTE program, e.g., the ratio of faculty to students, the number of alternative instructional strategies, etc.

New state certification standards which specify competency or performance by candidates also require that the new teacher education programs be managed by a consortium of agencies, including not only colleges or universities but also school administrators, professional associations, and the public. Released time for teachers to participate fully in training programs and in policy decisions is a new cost for teacher education.

The realistic level for funding CBTE might be to compare teacher education to graduate education or to other high-cost programs such as in the laboratory sciences. The rationale for shifting the funding from a level comparable to academic college instruction to that of funding graduate or laboratory science instruc-

The higher level of funding per student does not mean that the total cost of teacher education necessarily would be higher; preparation agencies could limit enrollment in their CBTE programs.

[illegible] is that CBTE is like other high-cost colleges in structure. CBTE is a highly individualized program and should be supported in the same way as other individualized programs which are committed to support a highly individualized approach.

Eleventh, the preparation of teachers under a new, private-based model will need substantial financing in the same way as teacher education programs are presently funded. Specific CBTE programs must carry over-ride credits. The credits must meet the quality standards other college credits must meet. The credits must be sufficient to provide for the maximum time of faculty and for support programs. The additional funding needed to implement the individualized CBTE programs must come through the existing system of higher education. Therefore, the key of CBTE seems to rest with the decision makers who allocate funds to colleges and universities, and specifically CBTE will depend upon a more generous formula than is now used to fund undergraduate teacher education.

Assessment of Teacher Performance: What is Involved? What is the Cost?*

by

[redacted]

When the term "performance-based" is used as a descriptor for teacher training and/or teacher certification, certain essential conditions imply. What the teacher does is as important as what he knows. Training for performance skill development requires equal or greater attention than knowledge acquisition. At the end of training, mastery is used as the evaluative criteria rather than a list of courses completed and grades received. Information about the teacher's performance is obtained throughout the training period and subsequently in the operating classroom in order to determine whether the specified teaching skills are being acquired and used. Student performance also is studied in order to determine whether teacher use of certain skills relates to the level and quality of student learning.

Within such a framework, assessment of teacher performance is multi-faceted. It is considered for many different reasons. Teachers need a wide variety of teacher and student behaviors, and involve a diversity of individuals and instruments. Inquiry into several of these facets of assessment is the purpose of this paper. The discussion centers around choices that must be made regarding assessment and the costs associated with particular options. It builds upon the information obtained by the Teacher Education Division at the Far West Laboratory for Educational Research and Development during 4 years of developing and testing performance-based teacher training materials.

Why Assess Teacher Performance

Assessment of teacher performance can be undertaken for a variety of purposes. Among these are the following:

- To determine the average type of performance that can be expected of graduates or trainees
- To identify the range of performance that can be expected of graduates or trainees
- To provide a measurement acceptable level of performance for graduates or trainees
- To determine whether an individual's level of performance is up to the level of others involved in training
- To establish that all or a certain number of trainees or graduates have achieved the desired level of performance
- To determine the teachers of teaching skills a measure of success on the job (i.e., teachers perform skill and those which do not, and to identify to support)
- To establish a minimum acceptable level of performance for graduates or trainees
- To determine whether an individual's level of performance is up to the level of others involved in training
- To establish that all or a certain number of trainees or graduates have achieved the desired level of performance
- To determine the range of performance that can be expected of graduates or trainees

Each type of data has its own unique and distinct purpose. Assessment is conducted to determine average teacher performance, identify how well teachers perform certain specific skills, or establish a minimum acceptable level of performance. Each data determination regarding teaching can be useful in a particular way and teachers may be involved in a particular way. The information obtained serves as a guide for policy decisions such as the selection of the content (skills and knowledge) to be included in a preservice and in-service training program and for the development of the skill and knowledge implementation (or certification). Control information of the type we use, and is documenting how an effective, pre-service and training program should be (i.e., the proposed number of or teachers needing training in a given skill and the geographic areas in which particular types of training are necessary).

On the other hand, in the case how the planning of individual teacher's mastering program and in working with a particular individual how to improve certain specified teaching skills, evaluation provides information regarding that person's performance. Knowing precisely where a person is in relation to the specific skill sought along with a judgment as to the quality of the skill should be available for each individual's effort. The specificity of the profile may exercise depending upon whether the purpose of skill use will be to assist in starting a training

---

*Revised paper, World Assembly Conference on Performance-Based Teacher Education, New Orleans, Louisiana, February 27, 1974.

[illegible]

[illegible]

[illegible]

[illegible] [illegible] [illegible] [

[illegible] 12

[illegible] Cognitive Archaeology

[illegible] Part II [illegible] Cognitive Archaeology

[illegible] Part III [illegible]

12 [illegible] Cognitive Archaeology [illegible] Studies in Archaeology of [illegible]

12 [illegible] Cognitive Archaeology [illegible] Studies in Archaeology of [illegible]

12 [illegible] Cognitive Archaeology [illegible] Studies in Archaeology of [illegible]

[illegible] [illegible] [illegible] [

Bibliography

Davis, O. L. and Tinsley, D. C. Cognitive objectives revealed by classroom questions asked by social studies teachers. Peabody Journal of Education, 1967, 45, 21-26.

Dodl, Norman, et al. "A Catalog of Teacher Competencies—A Working Document." Tallahassee: Florida State University, 1972.

Elementary Teacher Education Models, A Summary. Journal of Research and Development in Education, Vol. 2, No. 3. University of Georgia, Athens, Georgia, Spring 1969.

Gallagher, J. J. Expressive thought by gifted children in the classroom. Elementary English, 1965, 42, 559-568.

Guszak, F. J. Teacher questioning and reading. The Reading Teacher, 1967, 21, 227-234.

Herbert, John. A research base for accreditation of teacher preparation programs. Accreditation and Research Problems, J. L. Burdin and M. T. Reagan, eds. ERIC Clearinghouse on Teacher Education, Washington, D.C., 1971.

Houston, W. Robert. "Strategies and Resources for Developing a Competency-Based Teacher Education Program." University of Houston, New York State Education Department and Multi-State Consortium on Performance-Based Teacher Education, 1972.

Husek, T. R. and Sirotnik, Ken. Matrix sampling in educational research: an empirical investigation. Paper presented at the 1968 convention of the American Educational Research Association, Chicago, Illinois.

Joyce, Bruce R. Estimating costs of a competency orientation. Paper presented at Multi-State Consortium on Performance-Based Teacher Education, New Orleans, February 1973.

Joyce, Bruce R. and Weil, Marsha. "Models of Teaching." Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Lord, F. M. and Novick, M. R. "Statistical Theories of Mental Test Scores." Reading, Mass.: Addison-Wesley, 1968.

Rosenshine, B. "Teaching Behaviours and Student Achievement." London: National Foundation for Educational Research in England and Wales, 1971.

Silberman, C. E. "Crisis in the Classroom." New York: Random House, Inc., 1970.

Smith, B. O. Certification of educational personnel. University of South Florida, unpublished paper, 1971.

Turner, R. L. Relationships between Teachers for the Real World and the elementary models programmatic themes and mechanisms payoffs, mechanism and cost. In B. Rosner (ed.), "The Power of Competency-Based Teacher Education." Boston: Allyn and Bacon, Inc., 1972.

The Teachers Viewpoint

By

SANDRA FELDMAN

I have been asked to address my comments specifically to the concerns of State Education Agency personnel who have been and will be making significant policy decisions concerning performance education. I intend to do just that—and I hope you'll forgive me.

Although it doesn't seem to be, it should be a source of worry to state education departments that teachers are suspicious and wary of them, that they are not looked upon by the profession as allies, that they are not looked to as a resource for the solving of educational problems, that they are not thought of as providing educational leadership.

It should be food for thought, at least, that often the policymakers in state education departments, in the business of prescribing for the schools, find themselves on opposite sides of the barricades from teachers in legislative and other controversies on school matters, ranging from certification to teacher education to "alternative" school plans or performance contracting, or vouchers, or teacher evaluation, or professional practices acts, or questions of school structure and governance—or a host of other issues.

Why?

Why don't state education departments have a healthier respect for classroom teachers, for their accumulated experience and expertise, for their organizations as an expression of teacher concerns?

Why aren't teachers more significantly involved in the process of decision-making at state education department levels. (And by teachers I do not mean the state education department personnel who once were in the classroom and feel themselves able, therefore, to speak for teachers. Nor do I mean selected individual teachers, handpicked by school administrations. I mean teacher leaders selected democratically by the teachers themselves, through their organization.)

I would like to hear what you think the reasons are: I think they are manifold and I will tell you what I think first.

One reason, in my opinion, is an incomprehensible inability of education administration to accept the fact that times have changed, that teachers are organized, that they insist on speaking for themselves and on having a voice in the decisions which affect not just their working lives, but the education of the children they teach.

Insofar as your invitation to this conference of teacher representatives demonstrates some consciousness of that, I congratulate you—and I intend to take full advantage of this opportunity. But I do so with a certain skepticism; as a participant in many conferences such as this one, I seldom see the fruits of such exchanges in concrete application back home.

Another reason for a lack of real cooperation between teachers and their state education departments is a result, I think, of some state education department officials seeing themselves as in the vanguard of educational change and seeing teachers as defenders of the status quo. To teachers, this characterization is ludicrous. They do not see their demands for drastically reduced class size, for massively increased social, psychological, and health services for children, for the maintenance of high professional standards as "old hat." They do not see performance contracting, "accountability," legislation like the Stull Act, differentiated staffing, and merit pay schemes as revolutionary improvements, but as backward steps and attacks on the integrity of the profession. They are not impressed by paper-weighty master plans or glossy brochures advertising "bold new steps" which usually do nothing whatsoever to help them in the classroom.

There is yet another reason, in my opinion, for the "credibility gap" between teachers and state education departments; and this one I will give a little more time to, for it will lead me to the subject at hand. That is, I fear that many state education department officials in policymaking roles (1) are under great pressure to respond to the now famous "crisis of confidence" in the schools, without the knowledge and resources to do something real to end failure, and (2) share in the growing public attitude that teachers are responsible

BEST COPY AVAILABLE

for school failure and either don't know how to do something about it—or don't want to.

So, if we are to succeed in achieving the very worthy goal of this consortium—which is supposed to be "devoted to improving the quality of State Education personnel and their decision-making abilities," we have to first establish that one important, essential way of improving that policy decision making is to listen to the teacher organizations. Then we have to talk about the crisis in confidence that spurred the performance-based movement—why we have it and what we ought to do about it.

Sandra Feldman

Crisis in Confidence

The schools do not appear to be meeting the demands of a changed economy, of a job market which requires more and more skills and increasingly higher levels of training. Whether or not the schools in the past were the vehicle for upward mobility, for providing the path out of poverty, most people believe that they were—and expect them to continue that function. This problem of higher expectations is compounded by the social turbulence of the sixties and by the racial conflict which so often was most visible around school issues.

Second, despite the recent "revelation" by Christopher Jencks that schooling cannot end economic inequality among the adult population—and despite the fact that this was undoubtedly always true, that economic equality would have to be legislated as a matter of social policy—public schooling has undeniably produced a literate citizenry. Where in the past—and especially through the thirties, forties, and fifties—school teachers were among the most educated people in our society, even when they only had a high school or training school degree, today, the country is full of educated people. Thousands upon thousands of parents and citizens in white collar government employment, in advertising, in media jobs, accounting, computers, in a wide variety of nonteaching academic pursuits, feel that if they were not working at their present occupation, they could teach. Today, although teachers are still among the most educated, they do not have the status that comes with having knowledge that others do not have. While a holder of a Ph.D. in economics feels that a doctor's or a lawyer's skills are very special, even somewhat mysterious and intimidating, he does not feel that way about a teacher—even if the teacher also holds a Ph.D.

As a profession, education must, like other professions, assemble its experience and skill into a concrete, coherent body of knowledge as tangible as what exists in law or medicine.

That is why we must have performance-based teacher education. Not only will it make teacher education more relevant to classroom needs and thus improve education, but it will, if developed properly, provide us with that concrete knowledge without which we cannot much longer defend the public school system.

And that is why I believe, unshakeably, that developing a body of knowledge about the teaching-learning process ought to be the initial thrust and prime concern at present of the performance-based teacher education movement. We should be setting about doing what Fred MacDonald is proposing be done by a National Commission on Research and Development. We should be making a coordinated, long-term commitment to validating teacher competencies —not what is being done, which is a short-term commitment to listing them. We should be working at proving what teacher behavior, what teaching strategies, effect what learning and how.

We should not, as the New York State Education Department is doing, be telling teacher education institutions to "Poof! Change over to a performance-based teacher education program for certification approval purposes." At a meeting several months ago, to which representatives of teacher education institutions and the state teacher organization were invited, there were as

many different notions of what performance-based teacher education is as there were deans of education at the meeting. And the leadership in the state education department is not providing a model or a research design or any real guidance—unless you call requirements listed on a form for program approval, guidance. We will have as many different, and as many irrelevant and inadequate teacher education programs as we do right now—but New York State will "brag" about its performance-based teacher education program at conferences like these throughout the country.

Just recently, at a meeting of sophisticated chairmen of education departments of one of our largest teacher education institutions, I saw education department representatives of a college struggle with a paper description of what they hoped would be considered a performance-based program; and they had listed not teaching competencies, but behavioral objectives—which is fine to do, but from which, as Bunny Smith has emphasized over and over, you learn nothing about the teaching process.

I have grave doubts as to whether the states can autonomously develop meaningful performance-based programs—but I have no doubt that most teacher education institutions, without resources for research and without access to a large variety of classrooms, cannot do it.

If the performance-based teacher education movement is going to be meaningful, it cannot be handled as a public relations gesture by state education departments and tossed back to the colleges which have been floundering all along. Nor should the effort to develop a knowledge base of teacher competencies be tied to certification at this time, and certainly not to continuing certification schemes, because teachers will not stand for the destruction of tenure and job rights through renewable certification—especially when a review of performance is based on unvalidated lists of competencies.

If performance-based teacher education is going to be meaningful, it cannot be—as every education "innovation" usually is—doomed to success. In any other profession—take medicine—strategies are tried and tested and discarded if they don't work and the knowledge gained from systematic research into those strategies, whether they cure the disease or not, provides a continuous buildup of intelligence in the field. In education, we scurry to hide—or ignore—failure. That the research we've done so far is inconclusive, because sophisticated instruments to successfully interpret and measure teacher behavior are still in the early stages of development—and I know you've heard MacDonald, Rosenshine, and Schalock—should not be a source of embarrassment to us. It should provide us with the resolve to push that important work forward.

Instead, each state goes its own way, without a pooling and sharing—not just of knowledge, but of political power to get a commitment for funds to do the research we all know needs to be done. In New York State we have a very definite timetable. By—is it 1984?—we will have performance-based teacher education, so-called, and performance-based teacher certification throughout the State. But will we know any more about education or about teaching? I think not—unless the National Commission for Research and Development is ready to share their work with us—and they haven't got nearly enough money to do the work that has to be done. But that is the direction we should be taking.

What we should be doing—and actually the multistate consortium might well be an excellent vehicle for this—is coordinating a single effort, a national effort to develop a knowledge base. We should be collating the limited knowledge we now have, and we should be building a small number of teacher models based on available research and the opinion of experienced classroom educators. We should be developing in a broad cross-section of schools throughout the Nation a systematic assessment and data-gathering machinery which would enable us to compare and study teaching behavior and its effects on learning response in a prescribed variety of school environments. We should teach prospective teachers in the models created—not hundreds of different ones, but a few—so that we can control for effects, and we should have a research design built in so we can validate the competencies we have agreed on by studying those new teachers, on the job, where their education should continue during an internship-probationary period.

While I believe that the main thrust of performance-based teacher education should, in the beginning, be preservice, there must be involvement of experienced classroom teachers because they have a great deal of expertise to offer and because they will learn a great deal in the process of participating.

This past summer, The City University of New York came to UFT to seek cooperation in the development of competencies for teacher-training models. In a short time, we recruited over 300 interested teachers (who were paid a fee for their participation, as they should be). They provided a valuable contribution to the work, and they learned a great deal.

"I was forced to think out in a concrete way my role and skills as a teacher."

"I was forced to analyze myself and think of what really makes a good teacher."

"It is high time an emphasis was put on specific competencies rather than abstract theories."

"I came away more optimistic about the future of education."

(And school critics are always seeking a way to better teacher "attitudes".)

State education departments should be joining with the teacher organizations to demand that National Institute of Education and its funds be put at our disposal, that it be used as a resource for the profession, not in the form of hundreds of diverse grants awarded in a scattered way throughout the country, but in a massive, concentrated effort to develop the profession in the way I have just briefly described.

As state education department personnel, you should not be plunging forward blindly, as I think you are, because of the great public pressure to "do something." You should insist on your right and your need to base your decision making on substantive, proven, professional knowledge. Whether you do or don't, the teacher organizations will insist that you do; and if there is a struggle between those demanding quickie public relations solutions and the teachers who are demanding substance, you will be squeezed in the middle —unless you have already chosen sides—and the promise of performance-based teacher education may die in the process.

Some Substantive Issues And Political Considerations In Performance-Based Teacher Education

By

[illegible]*

What Have We Heard?

We've heard here over the past 4 days that

1. the state of the art is little advanced—in fact someone termed it "primitive",
2. there is little data on what makes for competence—as some put it—"We know too little about what's worth teaching",
3. the most difficult job will be defining competencies,
4. efficacy of assessment apparatus is the most critical issue and will be most difficult to develop,
5. the definitions of behaviors are, at this point, almost totally hypothetical,
6. we lack summative evaluation tools, i.e., there are rating forms for making global judgments, and there are research tools containing specific skill formulations—but nothing in between,
7. there is a complex interaction between the setting and performance, an interaction that affects performance in a major way, one that we haven't much taken into account, let alone analyzed,
8. the affective domain is being neglected in performance-based teacher education,
9. the tools for measurement in the social sciences, including teacher education, are in an infant stage; i.e., even with the sophistication developed in measurement on cognitive skills such as reading, we're still not able to diagnose causes of all reading difficulties, let alone prescribe behaviors that will correct them—so, it has been asked, how will we ever be able to measure attitudes, values, and the like?
10. there is a monumental problem in matching teaching styles to learning styles,
11. finally we've heard that measuring specific performances at various criterion levels will be costly.

There are a lot of other lesser problems we've heard about, and likely some additional major ones I've omitted.

What Does What We've Heard Say To Us?

One thing it might say is "forget it." My starting off with such a laundry list could cause you to believe that's where I am—"forget it," or that I'm launching into a defensive diatribe.

Actually, I believe that performance-based teacher education is a promising and viable concept and should be pursued.

My listing these concerns is for the purpose of putting a framework around what I'm about to suggest we do, and that is the following question:

What Are the Main Issues in Performance-Based Teacher Education That Researchers, Developers, Teachers, Professors, and Others Should Be Emphasizing?

In my judgment, a first line of activity should be around tools for evaluation. Performance-based teacher education will rise or fall on our ability to evaluate. That is, a first-order question is "How shall we determine that teacher candidate A is now ready to practice the profession and should be licensed to do so and that candidate X should 'go back to the drawing board'?"

Another way of saying it is that until we can measure performance in such a way that we are able to confidently identify minimal levels of performance to practice the profession, we will not have performance-based teacher education.

Some have said at this conference that prior to identifying minimal levels of performance there are the more difficult jobs of

* Dr. McKenna is professional associate, Instruction and Professional Development, National Education Association.

[illegible] goals and objectives for education;
[illegible] performance competencies re-
[illegible] goals and objectives.

I don't believe these to be the most difficult tasks and once it they were, we're farther along in accomplishing more than we are on the evaluation matrix.

For example, there are available:

1. some well-developed multi-dimensional systems;
2. several quite complete taxonomies;
3. the works of Mager, Kibler, Popham and others on getting goals and objectives into more readily manageable performance-based terms;
4. "The Florida Catalog of Competencies" and "Signs of Good Teaching" (the latter is less known but highly developed);
5. the Seven Cardinal Principles or The Purposes of Secondary Education;
6. the techniques of more recent writings reported at this conference by Kaoly Thomas of The City University of New York in which they were able to gain high consensus on which skills should be most valued;
7. the Northwest Regional Laboratory materials and those of the Teacher Corps.

I'm not that concerned about our ability to arrive at goals and objectives.

Some in this conference have argued that "instructional systems," as Bruce Joyce termed them, will be more difficult to develop. I'd argue that we have made quite satisfactory progress on those also.

Whether one goes with those toward the affective end of the continuum like "What a Course of Study" and the Taba Social Studies Curriculum or the more behavior-modification oriented types reflected in some of the packaged curriculums, there is a variety of quite well-developed instructional systems to choose from which include both content and teaching strategies.

So that brings us back to evaluation devices--and I repeat--performance-based teacher education will rise or fall on our ability to evaluate performance in relation to agreed-upon standards.

What research on evaluation leaves much to be desired, and I'm convinced that too little emphasis is being given to research and development in this key area in present efforts with performance-based teacher education.

There are places to start, and I'm surprised that more of the programs haven't attended to these promising possibilities. For example, the performances identified by Hamachek and Lois Bolton, both at Eastern Michigan, enthusiastic, for their classes were providing information on competencies with students and on their place in them.

There are also affective tools that should be considered whether or not they are those about relationships to learning outcomes. The premise here is that there are many processes which may be not least directly or particularly goals but at least contributory to achievement of the very goals which should be paramount and which are worthy of evaluation for their intrinsic value. One brief illustration of this. A recent study of all school dropouts reported by a district of the Intermediate high school population indicated that great numbers of students may be dropping out, albeit not in the conventional sense, because they have limited opportunity or experience during their school careers. If this is so, it is not inconceivable that the process of an education in the schools becomes, as much as possible, a determinant of the kind of environment as it is of practices in the greater society. Such can the improvement whether or not such processes can be deliberately developed or controlled to specific learning outcomes? Pre-evaluation systems for performance-based teacher education should incorporate such criteria.

I believe, as Bruce is [illegible] earlier in this conference, that it will not take [illegible] to develop valuable and relevant evaluation tools. I am pretty sure [illegible] that the big suggested [illegible] Before asked if it were possible to develop an accountability model for New York City, Bruce, then at Educational Testing Service, [illegible] that it could be done, but would be more difficult [illegible] and it is well known that Mr. Dobbin has [illegible] that project, absent anticipated resistance and substantial power.

Three years later Dr. Dale McDonald and his people have made some progress on the New York City accountability situation. I'm not sure that the development of their model is appreciably promising as first look. But it almost totally skirts the big problem of relating particular [illegible] to student achievement. What might be the cost and the time [illegible]?

Ted Kaye told me [illegible] Sunday that Educational Testing Service [illegible] years or so back ago that there are some 30 or 35 variables which do have marked performance which contribute to student or achievement. Even when put to good use recently when he was consultant [illegible] that "aside from the socio-economic [illegible] the schools of standardized achievements [illegible] teachers a major."

[illegible] [illegible] [illegible] [